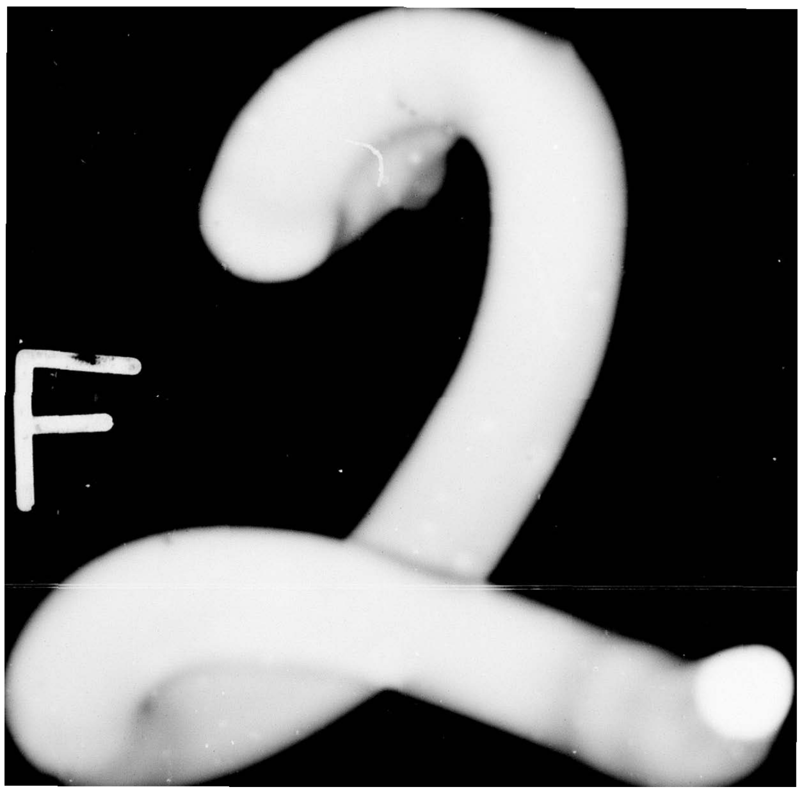
AD-A048 193 HUGHES AIRCRAFT CO CULVER CITY CALIF DISPLAY SYSTEMS--ETC F/G 17/9
SYNTHETIC APERTURE RADAR IMAGE PROCESSING TECHNIQUES DEVELOPMEN--ETC(U)
OCT 77 E J DRAGAVON, M L HERSHBERGER F33615-76-C-1115

UNCLASSIFIED HAC-P77-189 AFAL-TR-77-113 NL

1 OF 2

AD
A048193


F

AFAL-TR-77-113

AD-A048193

(1)

SYNTHETIC APERTURE RADAR IMAGE PROCESSING TECHNIQUES DEVELOPMENT AND EVALUATION

Display Systems Department
Hughes Aircraft Company
Centinela and Teale Streets
Culver City, California 90230

OCTOBER 1977

Final Report for Period March 1976 to February 1977

AIR FORCE AVIONICS LABORATORY
AIR FORCE WRIGHT AERONAUTICAL LABORATORIES
AIR FORCE SYSTEMS COMMAND
WRIGHT-PATTERSON AIR FORCE BASE, OHIO 45433

NOTICE

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely related Government procurement operation, the United States Government thereby incurs no responsibility nor any obligation whatsoever; and the fact that the government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

This report has been reviewed by the Information Office (I)) and is releasable to the National Technical Information Service (NTIS). At NTIS, it will be available to the general public, including foreign nations.

This technical report has been reviewed and is approved for publication.

Edmund G. Zelnio
EDMUND G. ZELNIO
Project Engineer

Gale D. Urban
GALE D. URBAN
Technical Manager
Development Group

FOR THE COMMANDER

Thomas E. Wenzel
THOMAS E. WENZEL, Maj, USAF
Act'g Deputy Chief
Radar Branch
Reconnaissance & Weapon
Delivery Division

Copies of this report should not be returned unless return is required by security considerations, contractual obligations, or notice on a specific document.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER AFAL-TR-77-113	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) SYNTHETIC APERTURE RADAR IMAGE PROCESSING TECHNIQUES DEVELOPMENT AND EVALUATION		5. TYPE OF REPORT & PERIOD COVERED Final Technical Report March 1976 to Feb. 1977 6. PERFORMING ORG. REPORT NUMBER D5676, P77-189
7. AUTHOR(s) E. J. Dragavon, M. L. Hershberger and P. J. Whitt		8. CONTRACT OR GRANT NUMBER(s) F33615-76-C-1115
9. PERFORMING ORGANIZATION NAME AND ADDRESS Hughes Aircraft Company; Display Systems Department; Centinela & Teale Streets; Culver City, Ca., 90230		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Project 7622, Task 01 Work Unit 19
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Avionics Laboratory (RWM-5) Wright Patterson Air Force Base, Ohio		12. REPORT DATE June 1977 13. NUMBER OF PAGES 118
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release: Distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

SYNTHETIC APERTURE RADAR, RADAR, IMAGE PROCESSING, IMAGE ENHANCEMENT, FEATURE ANALYSIS

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Synthetic aperture radar (SAR) has a dynamic range much greater than displays used to present the radar ground map video to human observers. This study program was undertaken to explore techniques by which the information contained within the large SAR dynamic range could be presented to human observers. In the pursuit of this goal, monochromatic adaptive gray shade transform, pseudocolor encoding, and feature analytic techniques were examined. Although some of the monochromatic and pseudocolor techniques resulted in improved quality of SAR video images, the improvement was judged

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. Abstract (Continued)

insufficient to justify implementation of such techniques in real- or near-real-time SAR systems. Feature analytic techniques used as automatic search, detection, and cueing aids to human observers showed considerable promise, and further work was recommended in this area.

PREFACE

This study was initiated by the Air Force Avionics Laboratory, Wright-Patterson Air Force Base, Ohio, to investigate enhanced synthetic aperture radar operator performance through better utilization of sensor dynamic range. The research was conducted by the Display Systems Department of Hughes Aircraft Company, Culver City, California, under USAF contract F33615-76-C-1115. The contract was initiated under Project 7622, Synthetic Aperture Radar Operator Performance, Task 01. Mr. F. P. Johnson (AFAL/RWM-5) was the Air Force Project Engineer during the early formative stages of the program; Mr. E. Zelnio (AFAL/RWM-5) took over as Air Force Project Engineer during the remaining phases of the program, Mr. M. L. Hershberger of Hughes Aircraft Company was Project Manager, and Mr. E. J. Dragavon was Project Engineer. The research sponsored by this contract was initiated March 1976 and completed February 1977. This report was submitted by the authors April 1977.

The authors gratefully acknowledge the assistance of Mr. P. W. Gregory of Hughes Aircraft Company for the computer programming and image processing of the monochromatic enhancement techniques.

CONTENTS

CONTENTS (Continued)

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS (Continued)

LIST OF ILLUSTRATIONS (Continued)

LIST OF TABLES

1.0 INTRODUCTION AND SUMMARY

INTRODUCTION

The continuing development of air-to-ground synthetic aperture radar systems provides significant potential for all-weather radar strike and intelligence collection. Whether this potential is exploited will depend on whether the performance capability of the radar can be utilized by the system operator in his effort to detect and recognize targets.

The target detection and recognition process can be viewed as an information channel consisting of several components. Each of these components shape overall system performance. The radar acts as an information source. If for some reason the radar produces signal information that is not passed along to the operator, the performance potential of the system is reduced. Thus, the effectiveness of the entire system will be constrained by the information transmission properties of limiting apertures in the system.

Synthetic aperture radars produce ground map image data with a radar signal return dynamic range that is typically in excess of 10,000,000 to 1. In comparison, conventional display media possess far less dynamic range, typically on the order of 1,000 to 1 for photographic film and 100 to 1 for cathode ray tube displays. The human visual system has an overall luminous intensity dynamic range of 10 billion to 1. However, instantaneous human visual dynamic range is on the order of 500 to 1. It is this instantaneous visual dynamic range that is important in real- and near-real-time sensor display systems.

The most effective use of synthetic aperture radar (SAR) may well be limited by 1) the inability of display devices to present the large SAR dynamic range, and 2) the restricted instantaneous dynamic range of the human visual system to perceive the large SAR dynamic range. Traditional approaches to this problem have used linear gamma compression, where one intensity level on the display is used to present a span of adjacent intensities in the incoming radar signal, or gamma shaping i.e. concentrating the entire presentation range of the display in a dynamic range region of the radar signal space. The first of these approaches results in a loss of detail in the presented image, while the latter disregards or highly compresses an entire region of signal intensities. Clearly, neither of these fixed gray shade transformation approaches fully exploits the capabilities of the sensor.

The objective of the work reported here was to explore methods of more purposeful assignment of incoming radar signal intensity information to limited display capability. A basic assumption was that the radar operator could use the information contained in the large radar dynamic range if it could be displayed in some manner. Evidence to support this assumption can be found in Humes, Craig, Poplawski, Guerin, and Hershberger, 1974; and Lamonica, 1975.

SUMMARY

Three avenues of investigation were pursued during the course of the study program. These were 1) monochromatic adaptive gray shade transformations, 2) pseudocolor encoding of SAR dynamic range, and 3) feature analytic techniques. The source SAR image data was 20-foot resolution, 4-look radar from the Forward-Looking Advanced Multimode Radar (FLAMR) system on magnetic tapes supplied by the Air Force. Computer processing was performed to investigate various SAR dynamic range enhancement techniques in each of the three areas of investigation. The techniques tried and the results obtained are described below. Although the investigation of feature analytic techniques was preliminary in nature, the results were all positive, and therefore these techniques show the most promise.

Monochromatic Enhancement Techniques

Eleven adaptive gray shade processing techniques were evaluated. Some of the techniques had been previously developed for other sensors, and some of the techniques were developed specifically for SAR during the program. Most of the enhancement techniques were one-dimensional (the SAR range image dimension). Table 1 lists the 11 adaptive gray shade techniques evaluated. Section II of this report describes in detail each of the 11 techniques and the results obtained with each.

TABLE 1. MONOCHROMATIC PROCESSING TECHNIQUES

Local area brightness and gain control
Local area histogram equalization
One-dimensional Haar Stockham nonlinear filter
One-dimensional Haar nonlinear crispening filter
One dimensional Haar linear high pass filter
One-dimensional Haar linear crispening filter
One-dimensional Fourier modified crispening filter
One-dimensional Fourier multilobe filter
One-dimensional Fourier multinotch filter
Two-dimensional Fourier modified crispening filter
Two-dimensional Fourier matched filter

Of the seven one-dimensional techniques evaluated, the Fourier modified crispening filter produced the best results. Figure 1 shows two FLAMR SAR images processed with the one-dimensional Fourier modified crispening filter. Comparing the original unenhanced images to the enhanced images, improved edge sharpness and contrast are apparent. Amplification of image noise that occurred with the unmodified crispening filter was also considerably reduced with the modified crispening filter.

Since the Fourier modified crispening filter gave the most improvement in SAR image quality of any of the one-dimensional techniques, the two-dimensional version of the Fourier modified crispening filter was implemented. Except for some intensity smearing problems that can be handled by thresholding prior to filtering, the two-dimensional filter produced increased enhancement effects. Figure 2 shows the two FLAMR images processed with the two-dimensional Fourier modified crispening filter.

Although the 11 monochromatic enhancement techniques had the intended effects on the SAR images, and generally the observed effects were desirable, the effects were judged to be of insufficient magnitude to yield any measurable improvement of operator performance. Such techniques are therefore not recommended for high resolution SAR tactical target acquisition systems.

1

2

Original

Enhanced

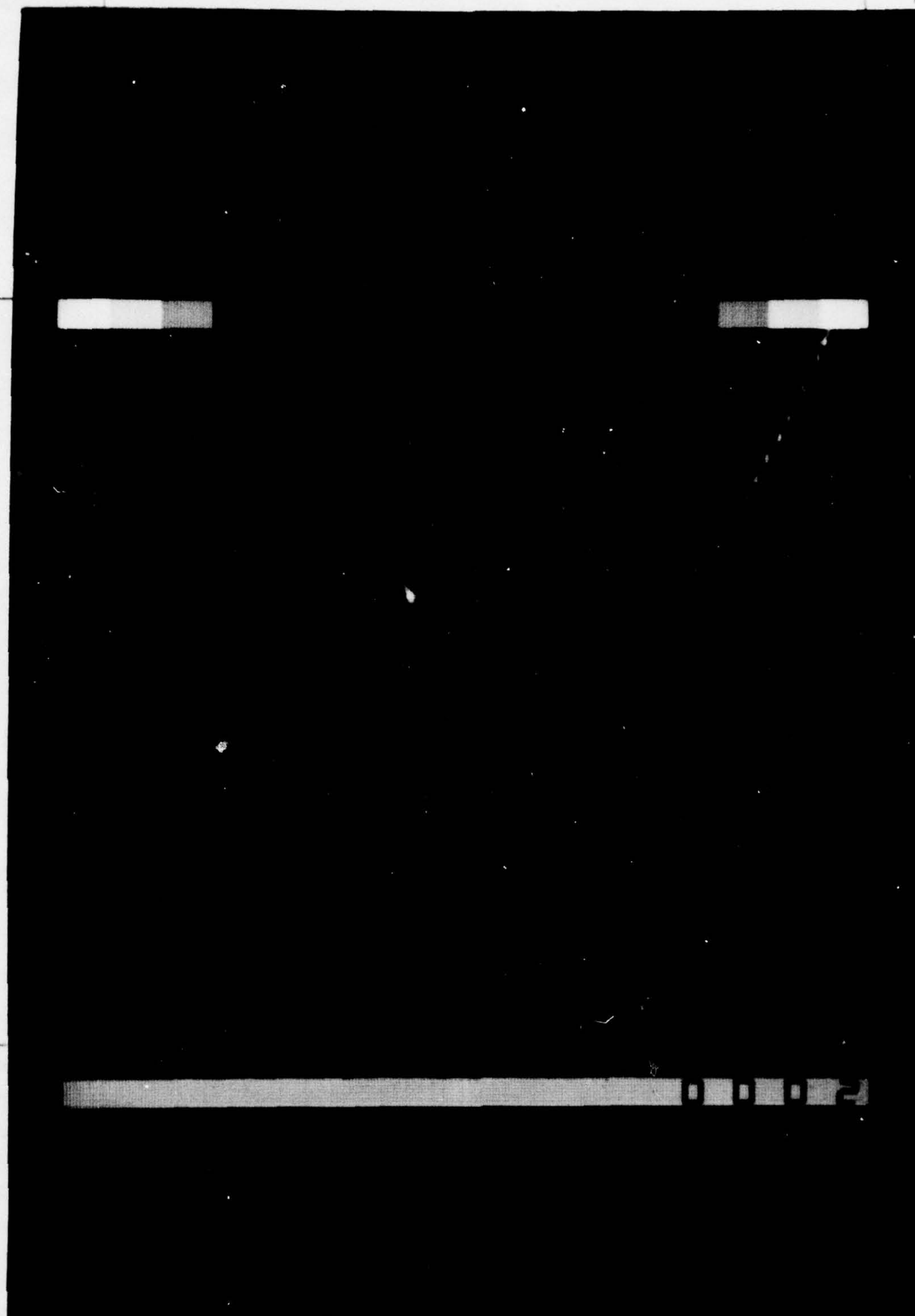
3

4

Original

Enhanced

Figure 1. One-Dimensional Fourier Modified Crispening Enhanced SAR Images



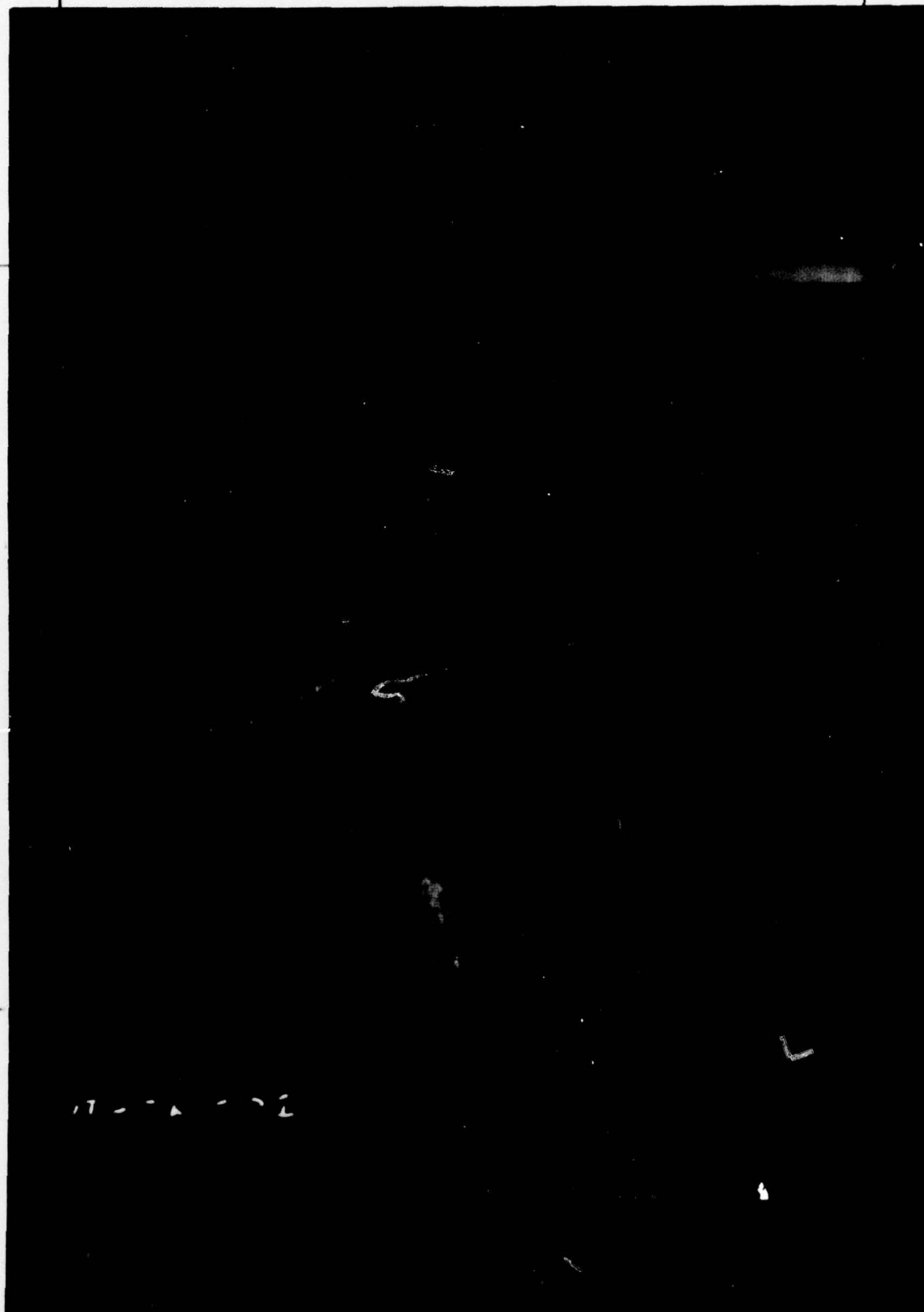
3"

to 79%

29a

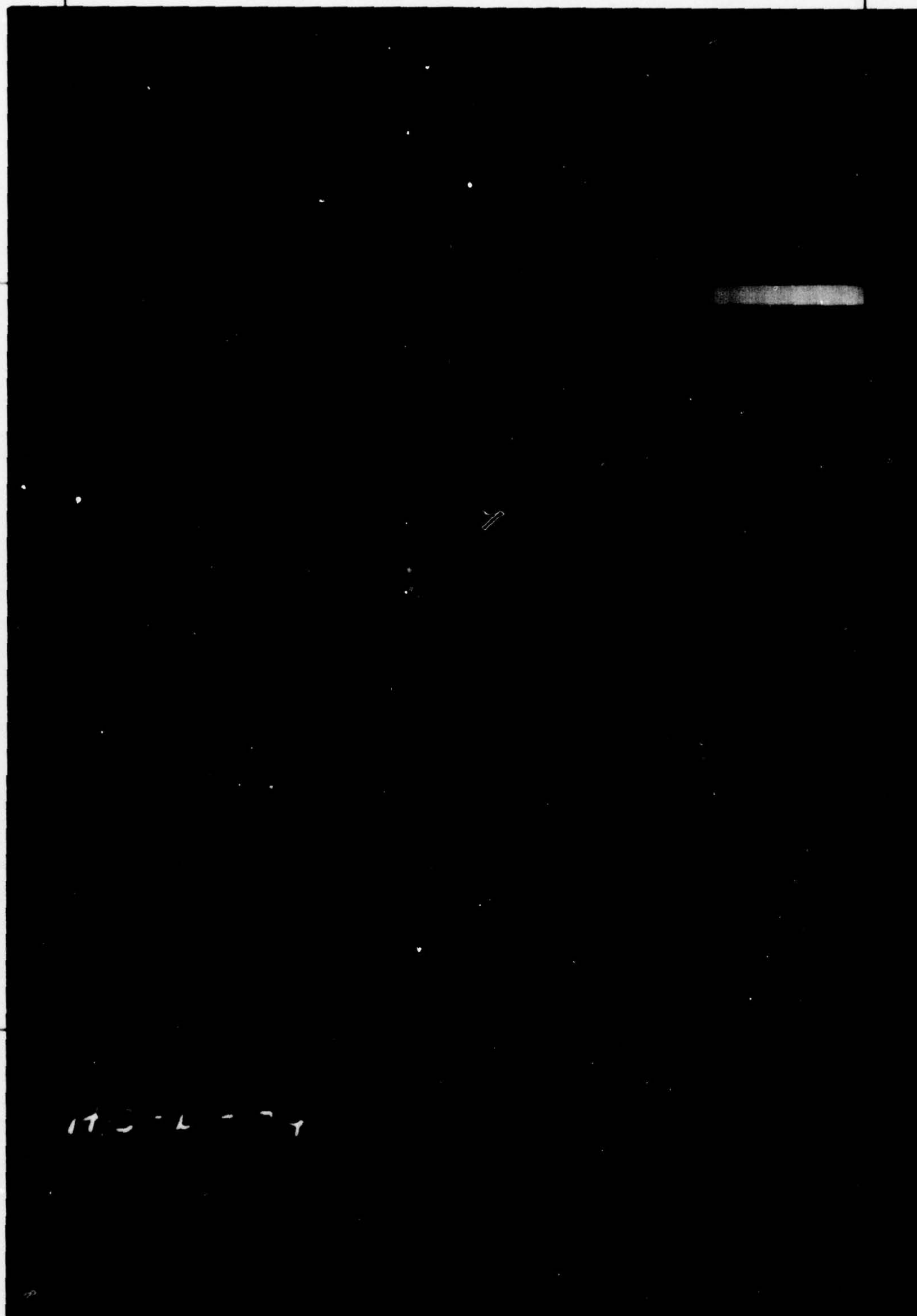
~~[illegible]~~ 25a, F. 1A, ~~[illegible]~~, ~~[illegible]~~, 10A

~~[illegible]~~, 27A, 13A, 16A, 19A, 22A, 23A

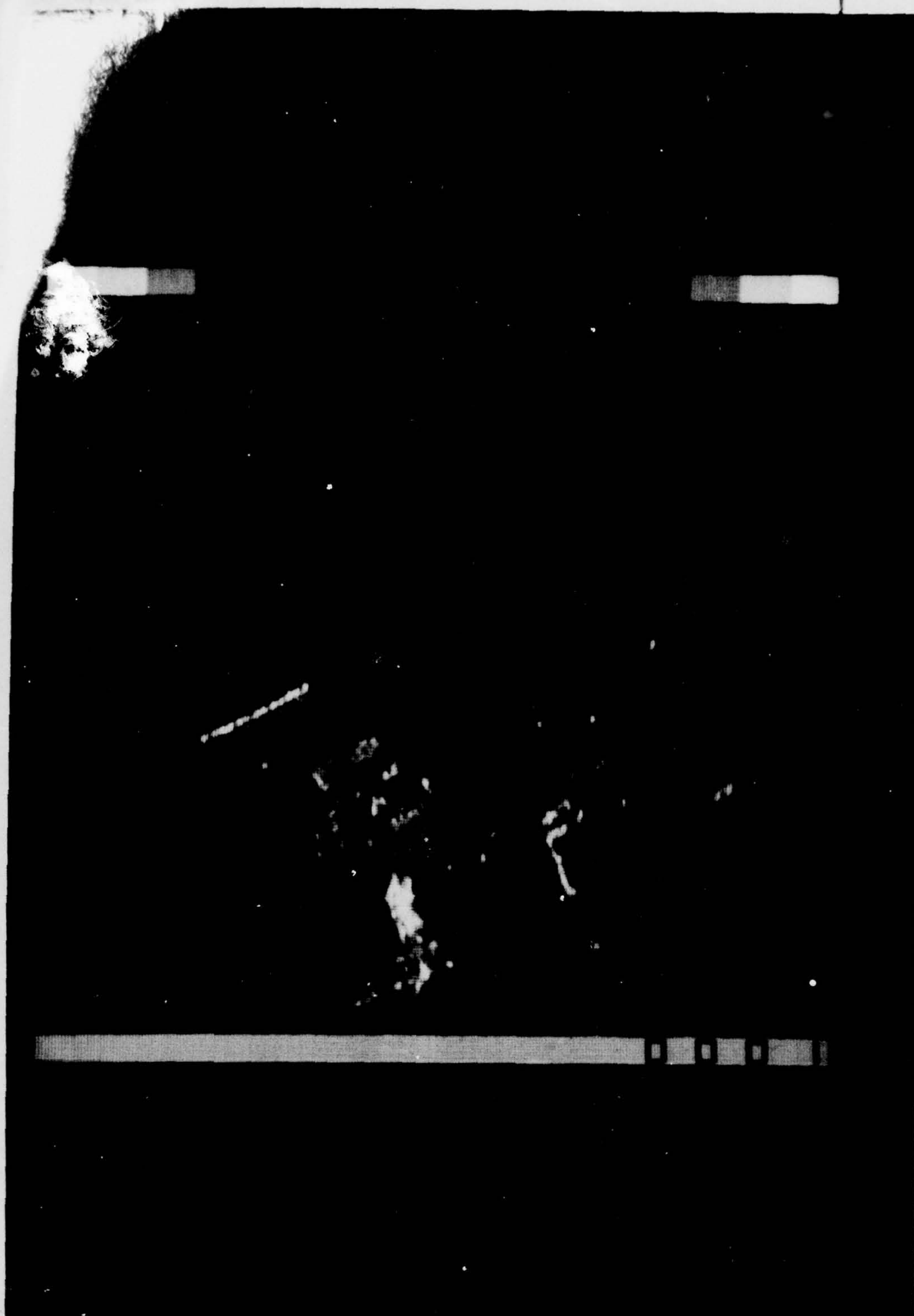


F. 1d ✓
25d ✓

&79%



F. 1B
25b to 79%



3"

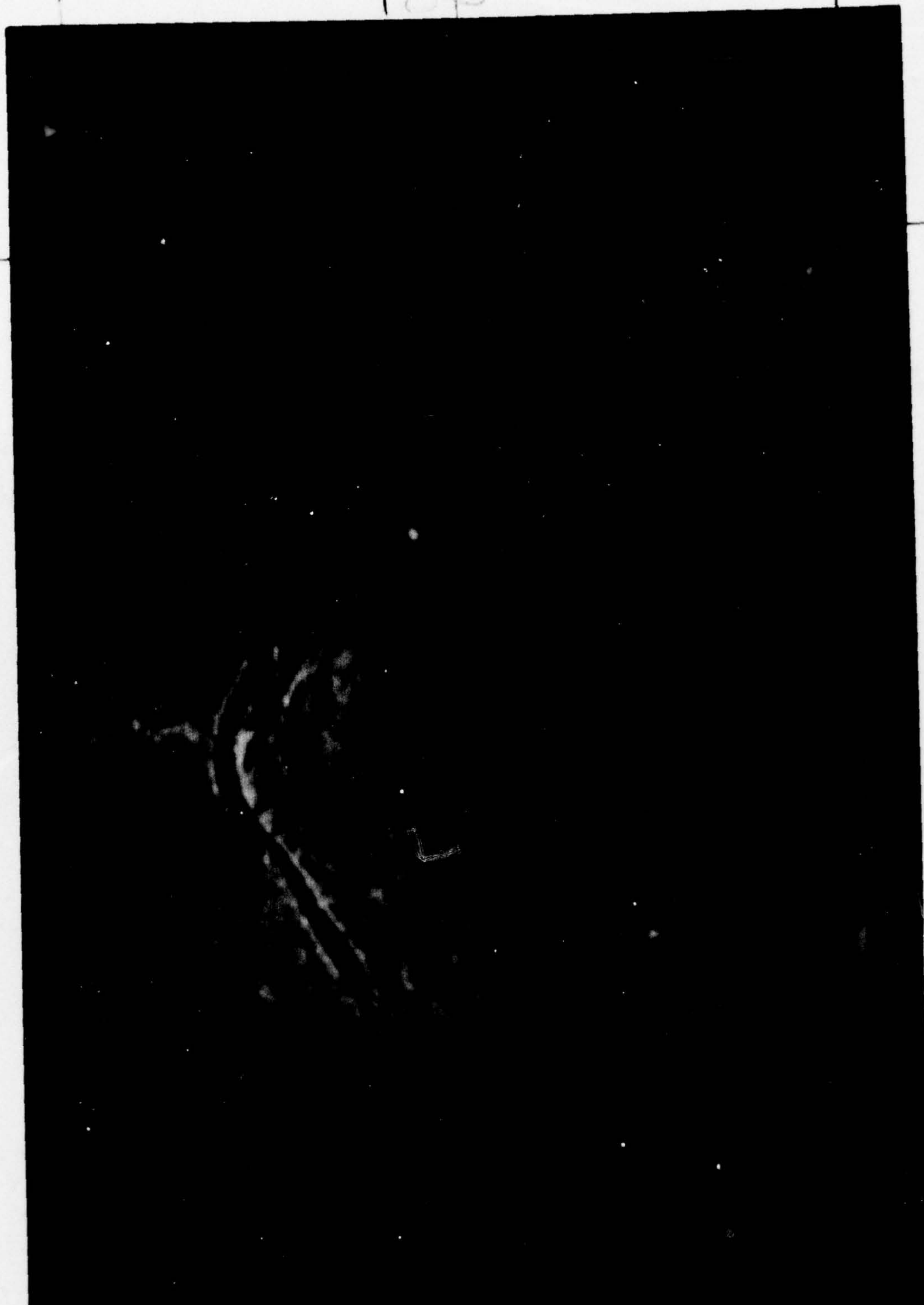
+079%

29C

~~41C, 38C~~ 25C, F.1C, [redacted], 10C

[redacted], 27C, 13C, 16C, 19C, 22C, 23C

TOP



F 2a, 32a, 33a

68%

5.8-5 (a)

to 79%

TOP

F. 2c, 32c + 33c

68%

5.8.5 (c)

to 79%

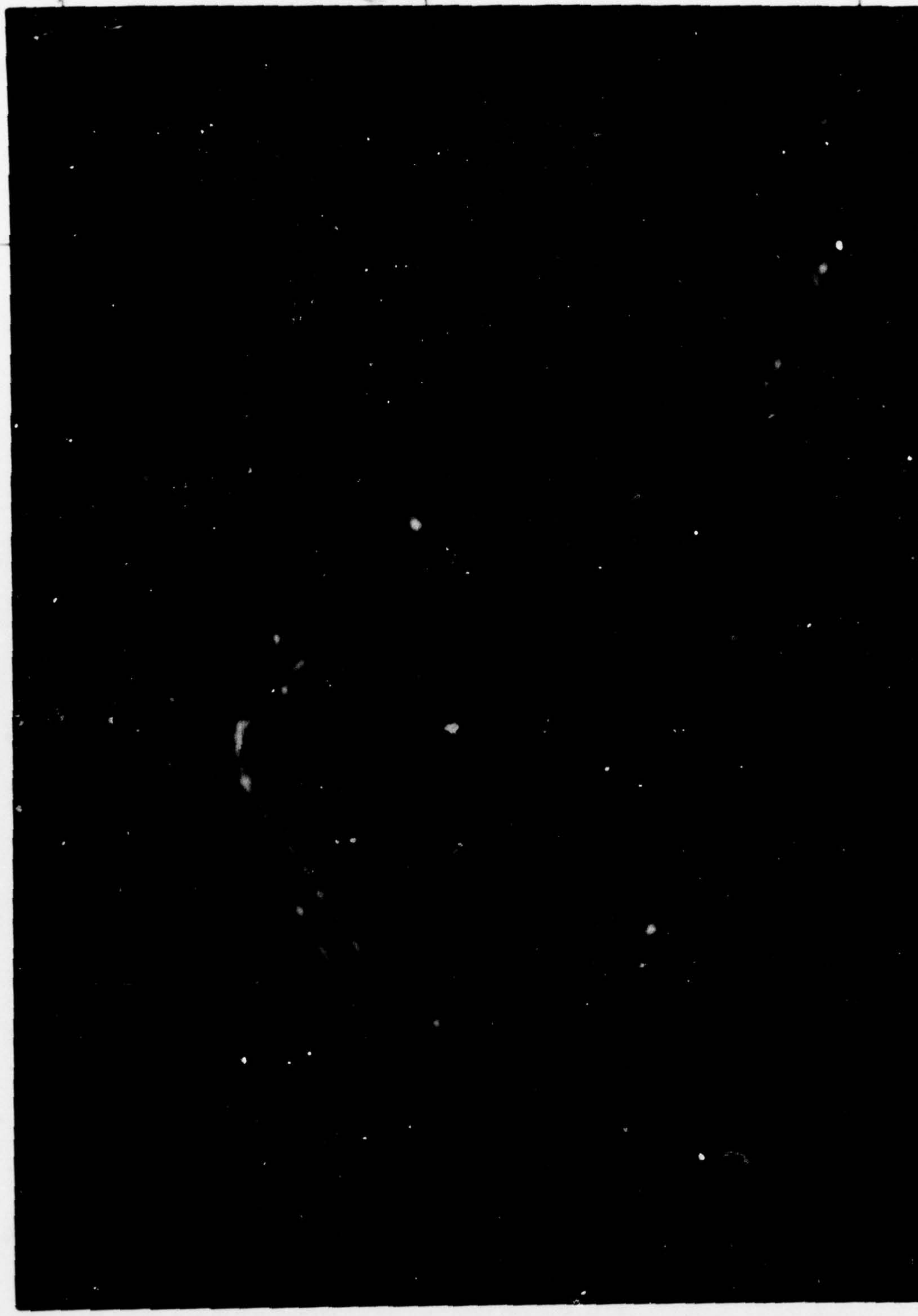
TOP

F.2d

~~30d~~

+07920

TOP



F. 2B
~~30 B~~

+079%

1

2

Original

Enhanced

3

4

Original

Enhanced

Figure 2. Two-Dimensional Fourier Modified Crispening Enhanced SAR Images

Pseudocolor Enhancement Techniques

SAR images possess no inherent color properties; hence, using color to display SAR images is referred to as pseudocolor encoding. Seven pseudocolor SAR encoding techniques were evaluated. The seven techniques, listed in Table 2, were placed into three categories: 1) one- and two-color intensity coding, 2) three-color intensity coding, and 3) hue coding and saturation coding. The use of color was investigated both independent and redundant to luminous intensity as a means of presenting large dynamic range SAR video.

TABLE 2. PSEUDOCOLOR ENCODING TECHNIQUES

One-color continuous intensity coding
Two-color discontinuous intensity coding
Two-color continuous intensity coding
Three-color discontinuous intensity coding
Three-color continuous intensity coding
Hue coding
Saturation coding

The color work was done using the Aerojet color CRT display and associated image memory, color map memory, and computer software at the University of Southern California Image Processing Laboratory. FLAMR SAR image data on magnetic tape again was used as the source radar maps. Appropriate distortion compensation techniques were employed in the generation of the color SAR images on the CRT and on color film. Section III of this report describes in detail each of pseudocolor encoding techniques and the results obtained.

Hue encoding produced the most positive results of the seven pseudocolor techniques evaluated. Two FLAMR SAR images processed by the hue encoding technique are shown in Figure 3. Besides being aesthetically pleasing, color differences appear to accentuate some of the features in the SAR scene, such as the river from its banks. It also appears that radar return intensity differences are easier to see in the color image.

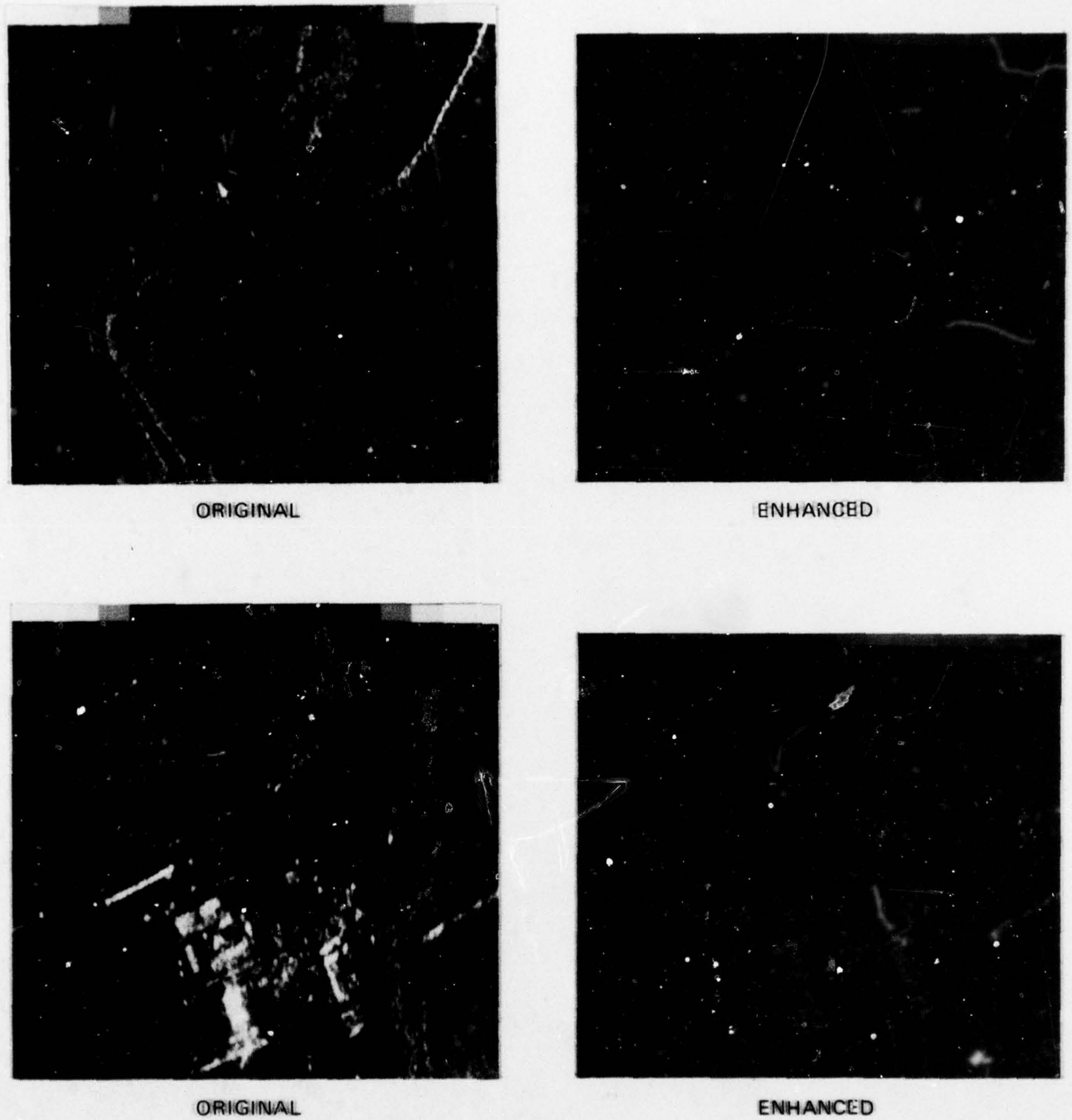


Figure 3. Hue encoded SAR images.

Although hue pseudocolor encoding seems to enhance SAR images, it is difficult to say how much operator performance improvement might be achieved with hue encoding. Our best estimate is that little improvement in operator tactical target acquisition performance would be realized. The best potential, we believe, lies in the area of detailed intelligence extraction.

Feature Analytic Techniques

The objective of feature analytic techniques is to search for properties in an image which are attributed to targets of a specific type. Automatic processing is done during the search process, and if enough properties of a given target are found, a human observer may be cued to the area on a displayed sensor image. Such feature analytic techniques applied to SAR could result in substantial reduction in tactical target acquisition time and increased probability of acquisition. An initial examination of SAR feature analytic techniques was the final task in this program. The key element in successful application of feature analytic techniques is the definition and selection of feature measurements to be made. This initial effort therefore concentrated on feature definition.

A scene analysis software package developed at Hughes provided the basic tools. Special SAR tactical target images were created by imbedding tactical targets (SAM site, AAA site, vehicle convoys) in a FLAMR background image. The mathematical feature analytic operators examined included: a statistics operator, edge detection operators, and a moment operator.

The size of the standard deviation and the distance between the local mean and the local minimum and maximum radar return intensity values were found to be useful statistical operators to distinguish image areas containing targets from image areas without targets. A Roberts edge detector operator was found to consistently detect targets even when the target-to-background contrast was reduced. The Roberts operator proved to be superior to the more complex Sobel and Hueckel operators. The moment operator, using the ratio of the moment about the horizontal to the moment about the vertical dimension of targets, provided data that distinguished the SAM and AAA sites from the convoys.

Based on the results obtained with the three types of feature analytic operators tested with synthetic aperture radar, the eventual reality of such an approach appears promising. The work performed in this program was only a first step, but the results were all positive. Although an operational SAR feature analyzer will require considerable development beyond what was done in this program, the potential payoff should be worth the investment.

2.0 MONOCHROMATIC ENHANCEMENT TECHNIQUES

BACKGROUND

The objective of this study was to investigate methods to achieve better utilization of the wide dynamic range of synthetic aperture radar (SAR) video. The class of black and white processing techniques discussed in this section were aimed at finding a process which would extract the most significant video intensity information and would present this information within a limited display dynamic range.

Digitally processed SAR video has a dynamic range much greater than can be instantaneously displayed on a monochromatic cathode ray tube (CRT) or a black and white film transparency. Thus, a linear mapping from radar return strengths to displayed brightness will be limited in one of the following ways: 1) potentially meaningful increments in radar signal strength are not perceivable on the display, or 2) lower or upper bound thresholding must be applied to the radar dynamic range. In either case, the full radar dynamic range is not displayed in a manner that the observer can utilize it.

Taking the logarithm of the radar return strength prior to a linear mapping to displayed brightness values will alleviate the thresholding problem somewhat. A penalty for taking the logarithm is that the range of return strengths corresponding to bright shades of gray is much larger than that corresponding to dark shades of gray, so that discrimination of small absolute variations in radar return strength is not uniform across the entire range of return strengths.

A logarithmic radar return strength to display brightness mapping with variable lower threshold and gray shade per decibels slope, as illustrated in Figure 4, was implemented in the Forward Looking Advanced Multimode Radar (FLAMR) system. Two FLAMR SAR images are shown in Figure 5. The FLAMR radar was the source of the synthetic aperture radar (SAR) digital image data used in this program.

As a part of the FLAMR program, experiments were performed which investigated the effects of variations in the lower threshold and the gray shade slope on SAR image quality. Experience with FLAMR and other

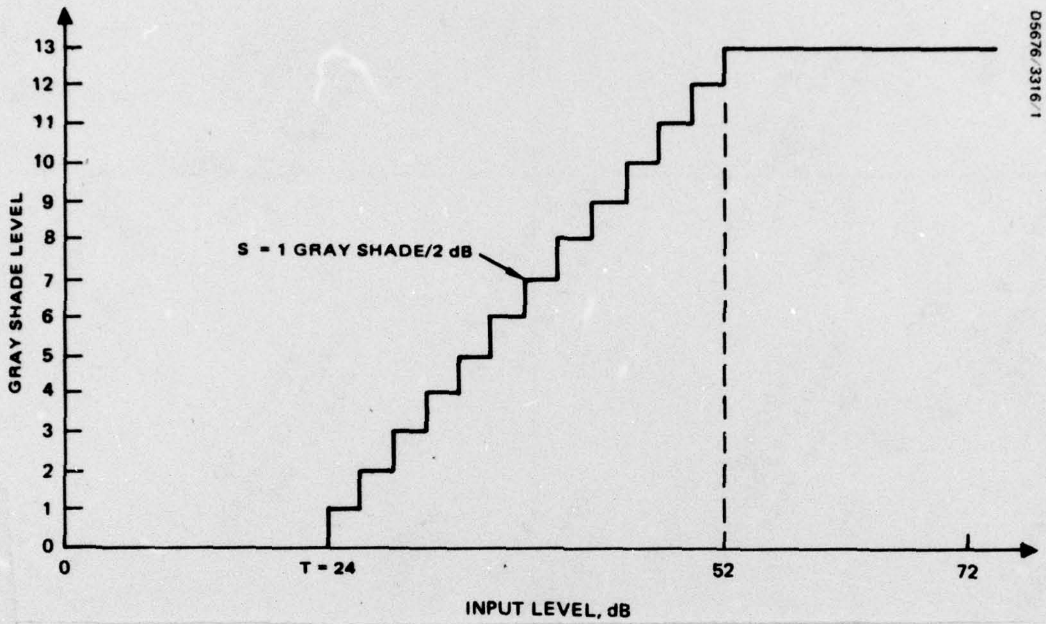


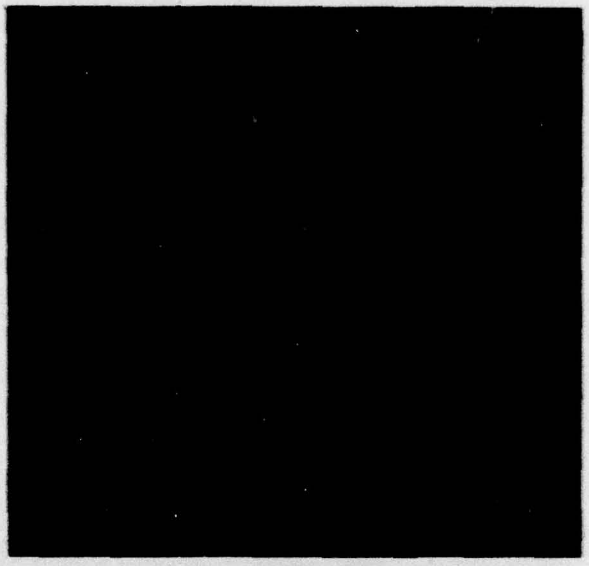
Figure 4. FLAMR gamma shaping function.

radar systems indicates that simple gray shade transformations offer limited potential for optimum utilization of SAR dynamic range.

APPROACH

Fixed gray shade transformations as a class have the disadvantage of not taking into account the nature of the scene being imaged and the characteristics of the target(s) of interest. In this program, selected existing and new adaptive gray shade processing techniques were investigated. The approach was to implement each processing technique in non-real-time on a general purpose digital computer. Original FLAMR SAR images supplied in digital form on computer compatible magnetic tape by the Air Force were processed according to the desired algorithm. The processed images were then written on a magnetic tape, and each image was digitally recorded on a film transparency.

The original images were either FLAMR scenes processed with the gray shade mapping shown in Figure 4, or the digital image data with preprocessing applied as described below. The latter originals were generated so that any differences due to different film types and/or different film processing would be minimized.



a. Bridge scene

b. Train scene

Figure 5. Example SAR scenes processed with standard FLAMR mapping algorithm.

Preprocessing

Twenty-one SAR images of various types of terrain from a FLAMR flight test program were supplied by the Air Force. The SAR image data was 20-foot resolution with 4:1 overlay (multiple looks). Table 3 lists the set of SAR images which were made available in digital form. The digital data were in the form of arrays of eight filter bins by 384 range bins, arranged sequentially on the magnetic tape as follows: eight doppler filter

TABLE 3. FLAMR IMAGE DATA

Image No.	Flight	Frame No.	7-Track Computer Tape No.	File No.	Scan Direction	Number of Arrays
201	21	024	4027	1	R-L	176
202	21	072	4027	2	R-L	177
203	21	215	4027	3	L-R	176
204	24	012	4027	4	R-L	176
205	24	034	4027	5	L-R	176
206	24	041	4027	6	R-L	176
207	24	057	4054	1	R-L	176
208	24	075	4054	2	R-L	176
209	25	034	4054	3	L-R	176
210	25	061	4054	4	R-L	176
211	26	452	4054	5	R-L	176
212	27	113	4054	6	R-L	176
213	27	062	4054	7	R-L	176
214	26	257	4054	8	R-L	176
215	24	047	4055	1	R-L	176
216	25	024	4055	2	R-L	176
217	27	066	4055	3	R-L	172
218	27	065	4055	4	L-R	176
219	34	120	4055	5	R-L	176
220	43	102	4055	6	L-R	176
221	39	015	4055	7	L-R	176

4

outputs corresponding to the first range bin, then eight filter outputs corresponding to the second range bin, and so on. After the eight filter outputs which corresponded to the 384th range bin, the next array of 8 by 384 points began. Each data point was a 24-bit word, of which only the 8 least significant (rightmost) bits were used. The 8 data bits were a binary number which represented eight times the logarithm to base 2 of the magnitude of the output of the doppler filter.

The preprocessing was accomplished as follows. The log filter magnitude data were read into the computer from magnetic tape, one array at a time. Each data array was decoded and stored into a two-dimensional array. Following the decoding, the data were raised to the appropriate power of 2 to convert to filter magnitudes. The data were then ready to be overlayed. The decision to perform the overlay in the actual filter magnitude domain rather than in the log filter magnitude domain was reached as follows. The FLAMR processor normally performed the log operation on the filter magnitude data and the overlay in the log filter magnitude domain. If a contrast enhancement or dynamic range utilization technique could be found which would effectively use the dynamic range of the actual filter magnitudes, it might be possible to avoid the log operation. On the FLAMR program, the overlay has been performed in both domains, and little difference exists for most scenes.

The next array of log filter magnitude data was read from the magnetic tape, decoded, and stored into a second array. The data in the array were also converted to filter magnitude by exponentiating. The data in the first array for doppler filters 1 and 2 were transferred to temporary storage arrays, and filters 3 through 8 were transferred one at a time to positions 1 through 6, respectively. Doppler filters 7 and 8 in the first array were then zeroed out. The filter magnitude data from the second array were then added to the corresponding shifted data in the first array. The data in the temporary arrays were then output to the disk for storage. Following this step, another array of log filter magnitude data was read into the second array, and the process was repeated until all data arrays in the image file on the tape had been read. At this point, the data in first array filter positions 1 and 2 were output to disk storage.

5

This process achieved the desired 4:1 overlay with 354 image lines of 384 range bins each created in a disk file for each image file of 177 filter arrays on the magnetic tape. Actually the first six image lines were not fully overlayed and were deleted; the resulting original image size on the disk was therefore 348 lines of 384 pictures elements per line.

The data in the disk files were represented as 32-bit floating point real numbers. Since fast Fourier and Haar transform techniques require data arrays whose length is a positive integral power of 2, the processing program was designed to operate on a 256 by 256 sub-array of picture elements (pixels) contigously located anywhere within the original 348 by 384 image data array.

Postprocessing

The original SAR image data were stored on a disk in floating point format to obtain maximum precision in subsequent processing without the need to change data formats during each processing run. Image processing techniques implemented at Hughes prior to the present contract, however, were set up to accept fixed point image data and could not easily be modified to accept floating point data. In either case there was a need to perform a floating point to fixed point data format conversion with 8 bits employed in the fixed point format. In the case of the prior processing techniques, this was necessary before they could be applied to the SAR data. In the case of the newly developed techniques, conversion was necessary after processing to display the processed image on film, because the film recording program accepted only 8-bit fixed point data.

The format conversion process referred to as postprocessing is an optimally scaled signed logarithmic mapping. The image array to be converted is first scanned by the computer, and four quantities are determined. The largest value of all the image data points is denoted by FMAX; the algebraically smallest value of all the image data points is denoted by FMIN. The smallest positive value of all the image pixels is FMINP; and the algebraically largest negative value of all the image pixels is FMAXN. Some processed images had negative pixel values, because the filter characteristic employed removed significant amounts of the dc, or average, component of the image. Because of the large dynamic range of the processed pixel values, a simple biasing of the processed pixels by the magnitude of the

6

algebraically smallest pixel value to insure pixel positivity prior to a logarithm operation would not suffice. This was due to dynamic range limitations of the floating point representation in the Sigma 5 computer used for processing. If an image had no negative pixels, FMINP was equal to FMIN, and FMAXN was set equal to -10^7.

A scale factor FSCALE was then defined to be the minimum of FMINP and the absolute value of FMAXN, i.e., FSCALE was the distance from zero to the nearest of FMINP and FMAXN. The positive logarithmic range of values, denoted by RANGEP, is then $\log_{10} \frac{FMAX}{FSCALE}$, and the negative logarithmic range of values, denoted by RANGEN, is $\log_{10} \frac{RMIN}{FSCALE}$, where RMIN is the absolute value of FMIN. The total logarithmic range of values, denoted RANGE, is the sum of RANGEP and RANGEN. RANGEP and RANGEN were always positive quantities, and RANGEN was zero if the image had no negative pixels, since in that case FSCALE was equal to FMIN and FMINP.

The quantity RANGE was then divided by 255, and this defined a quantity SLOT. Dividing RANGEN by SLOT and rounding to the nearest integer produced a quantity called KBIAS, which was the fixed point value which would be assigned to an image pixel of value zero. By construction, however, no image pixels had values between -FSCALE and +FSCALE. The signed logarithmic mapping then proceeded as follows. Assume a given image pixel has value V. If V was less than or equal to -FSCALE, its mapped value M was given by

$$M = KBIAS - \frac{\log_{10} \frac{-V}{FSCALE}}{SLOT}$$

rounded to the nearest integer. If V was greater than FSCALE, then M was given by

$$M = KBIAS + \frac{\log_{10} \frac{V}{FSCALE}}{SLOT}$$

again rounded to the nearest integer. The mapped values M were in the range from zero to 255.

With the foregoing understanding of preprocessing and postprocessing operations, we can describe the various monochromatic processing techniques evaluated.

IMAGE PROCESSING TECHNIQUES EVALUATED

Based on our inference that successful SAR gray shade enhancement techniques must be selective or adaptive in some sense, all of the monochromatic processing techniques which we examined possessed some form of image dependence property. The form of image dependence ranged from adaptation based on local image statistics to enhancement of selected spatial frequencies associated with a particular anticipated target size. Table 4 lists the monochromatic (black and white) processing concepts which were examined.

TABLE 4. MONOCHROMATIC PROCESSING TECHNIQUES

Local Area Brightness and Gain Control
Local Area Histogram Equalization
One-Dimensional Haar Stockham Nonlinear Filter
One-Dimensional Haar Nonlinear Crispening Filter
One-Dimensional Haar Linear High Pass Filter
One-Dimensional Haar Linear Crispening Filter
One-Dimensional Fourier Modified Crispening Filter
One-Dimensional Fourier Multilobe Filter
One-Dimensional Fourier Multinotch Filter
Two-Dimensional Fourier Modified Crispening Filter
Two-Dimensional Fourier Matched Filter

Local Area Brightness and Gain Control

The method of local area brightness and gain control (LABGC) is based on the fact that target-to-background contrast has a significant effect on detection probability and time to detect. The additional observation that background characteristics in the immediate vicinity of the target are more important than global background properties, particularly for small tactical targets, leads to the desire to improve local target-to-background contrast. This is the intent of the LABGC technique.

Local area brightness and gain control is achieved by sliding a small "window" through the image and adjusting the gain and brightness of the center element of the window. Figure 6 depicts this window which was chosen to be 9 pixels by 9 lines. If the gain function applied to the window

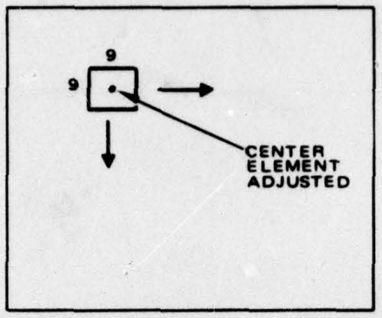


Figure 6. Sliding window for local area brightness and gain control.

center element is inversely proportional to the local standard deviation, as illustrated in Figure 7, the desired effect is achieved. One way to apply this gain is about the local statistical mean of the video. This causes a problem, however, in that certain pixels may exceed the bounds of maximum black or maximum white. This problem can be overcome by applying brightness control to bring the pixel intensities back within bounds. Figure 8b illustrates the application of video gain about the local mean to improve contrast. Figure 8c shows the application of the brightness control to avoid saturation. The combination of a gain function based on the local standard deviation and a brightness bias control comprise the LABGC contrast enhancement technique.

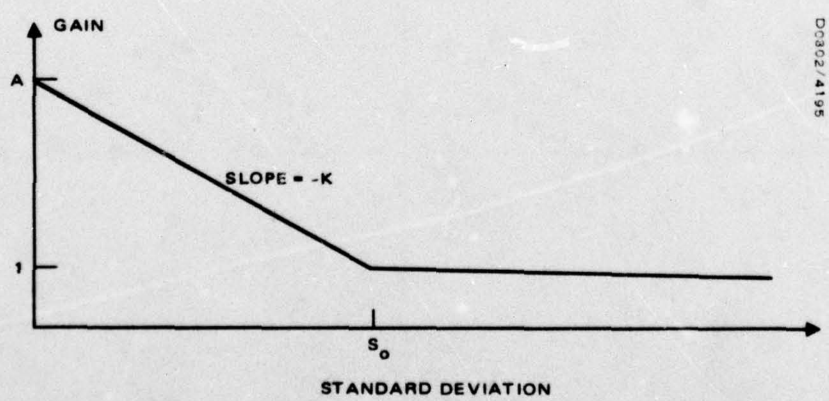


Figure 7. Gain as a function of standard deviation within sliding window.

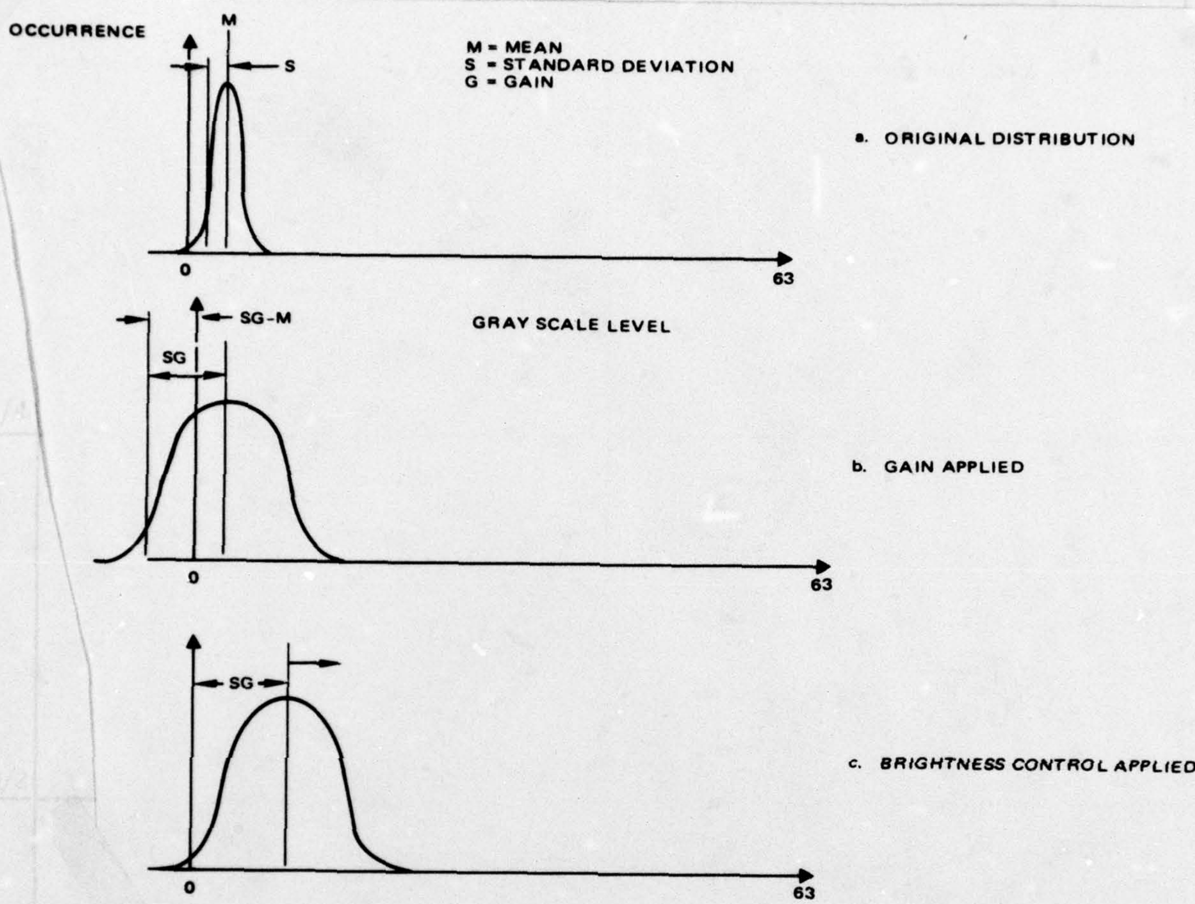


Figure 8. Statistical gain and brightness control.

Within the 9 by 9 sliding window the mean and standard deviation are computed as

$$\text{MEAN} = \text{M} = \frac{\sum_{i=1}^{9} \sum_{j=1}^{9} I(i, j)}{81}$$

and

$$\text{STANDARD DEVIATION} = \text{S} = \sqrt{\frac{\sum_{i=1}^{9} \sum_{j=1}^{9} I^2(i, j)}{81} - \text{MEAN}^2},$$

10

where $I(i, j)$ is a pixel intensity within the sliding window. The gain function is given by

$$\text{GAIN} = G = A - KS,$$

where A is the upper limit of the gain, K is the slope of the function, and S is the standard deviation within the window. If $G < 1$, then G is set to 1. The value of the brightness bias control is the distance that the video distribution extends beyond maximum black or maximum white, whichever happens to be applicable.

There are a number of parameters associated with the LABGC technique; among them are horizontal and vertical window size, maximum gain, cutoff standard deviation, shape of the gain function, and so forth. The effects of these and other parameters on the degree of contrast enhancement have been studied. The overall conclusion is that the technique is not sensitive to variations, within reasonable limits, of parameter values. Thus the results of LABGC processing of SAR images reported here are representative of the capabilities of the technique.

Since the computer program which performed LABGC in non-real-time already existed and required the input image to be in a 6-bit per pixel format, the logarithmic postprocessing technique described earlier was first applied to the original images. This produced image data with 8 bits per pixel. Figure 9 shows two scenes (bridge scene and train scene) with the 8-bit logarithmic postprocessing applied. These are the original unenhanced images against which the enhancement techniques will be compared. The 6 most significant bits of the 8 bits in these images were used as input to the LABGC program. Its output, also at 6 bits per pixel, was multiplied by four, so as to occupy 8 bits per pixel and therefore be comparable in potential dynamic range to the originals in Figure 9. The scenes resulting from LABGC F-9
processing are shown in Figure 10. In this section, the figures which show F-10
the enhanced SAR images contain four images: the bridge scene and the train scene, both unenhanced and enhanced.

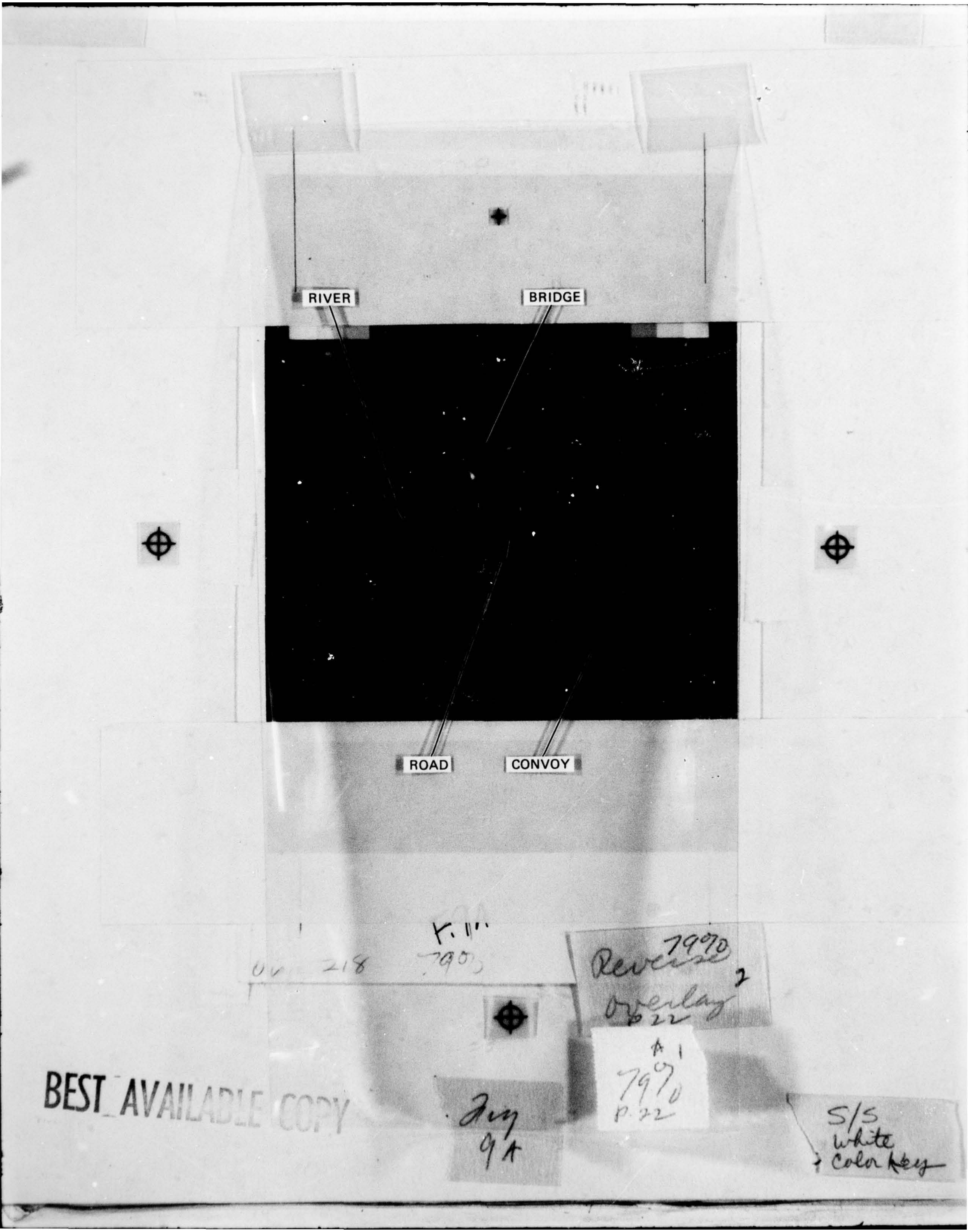
The LABGC technique was successful in improving target-to-background contrast in local areas. This can be seen along the edges of the river in both scenes. However, the land areas which appear fairly uniform in intensity in the unenhanced images have a more grainy character, and several areas, particularly in the bridge scene, appear to contain potential

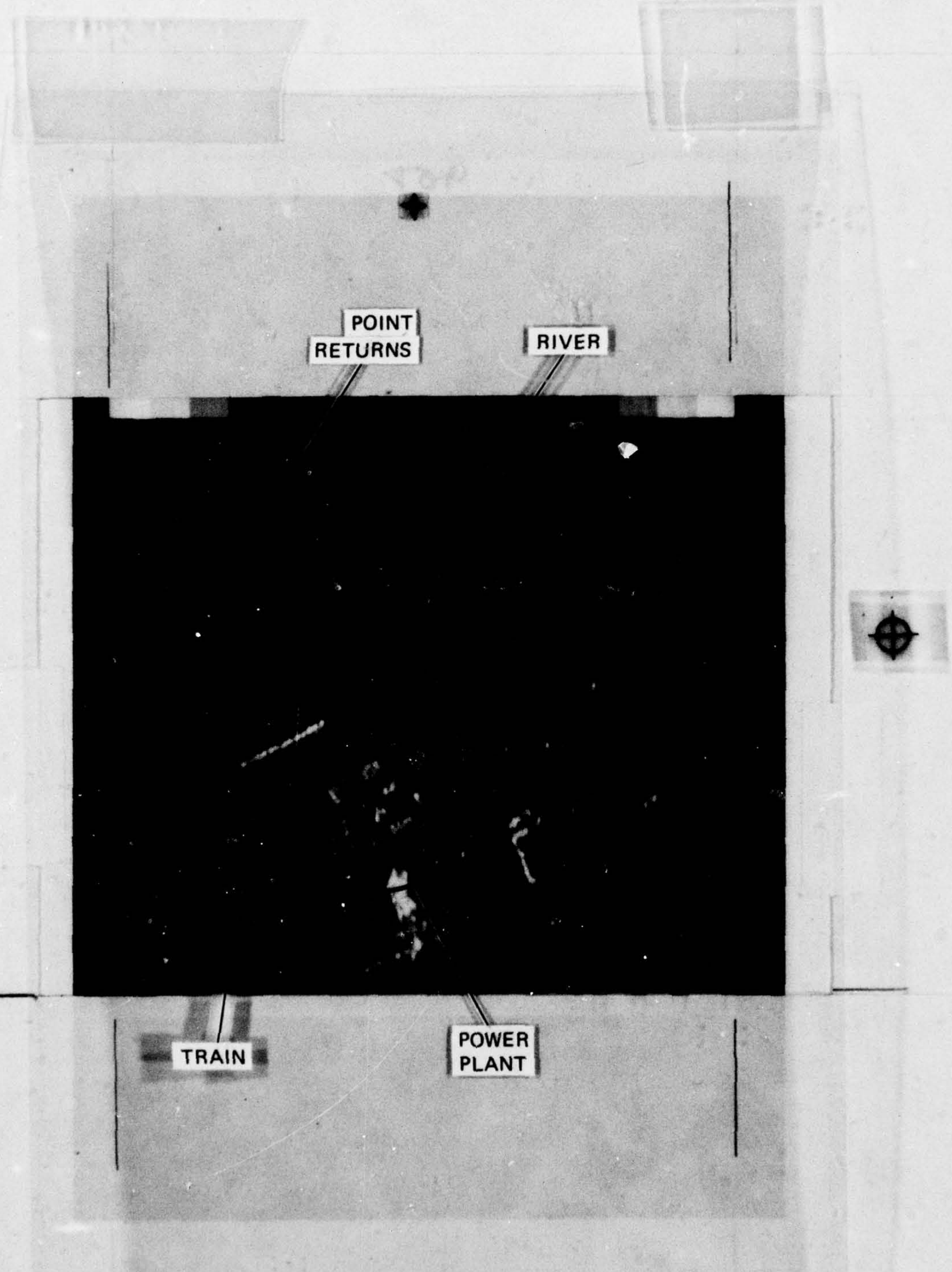
a. Bridge scene.

b. Train scene.

Figure 9. Eight-bit logarithmically mapped original SAR images.

12


RIVER
BRIDGE
ROAD
CONVOY
Reverse 7970
Overlay 2
p.22
A 1
7970
p.22
Fig
9A
S/S
White
Color Key



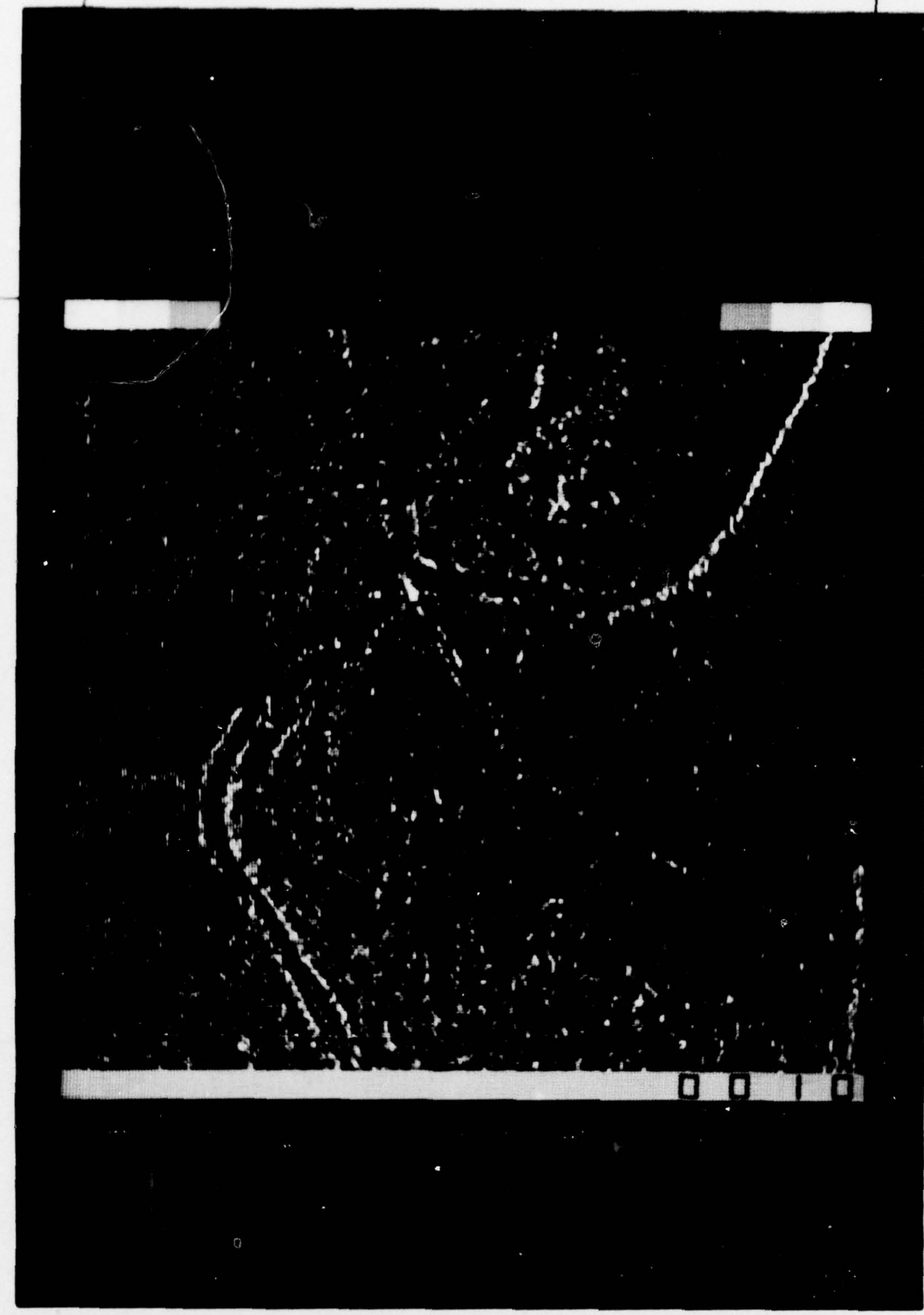
F.9B

rig 205 3

Reverse overlay 2

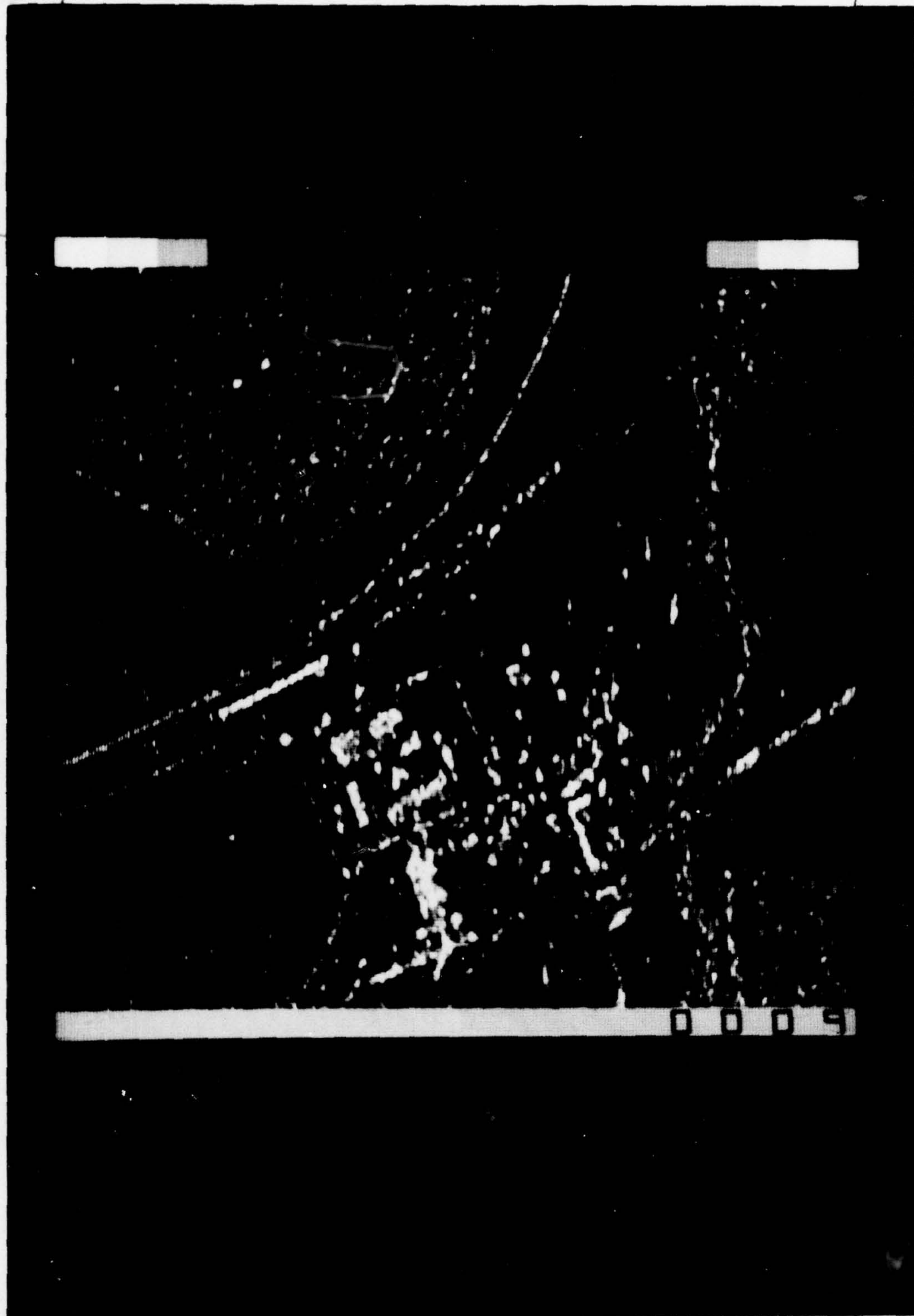
79% P22 1

79% P22



F10b

T079%0



F. 10d

J0792

1 2

Original Enhanced

3 4

Original Enhanced

Figure 10. Local area brightness and gain control enhancement.

1 13

small targets where there were none in the unenhanced image. This clutter amplification effect would probably negate any potential target detection performance improvement from contrast enhancement by requiring the operator to consider and decide upon a greater number of non-target objects.

Local Area Histogram Equalization

Local area histogram equalization (LAHE) enhances the relative contribution of small target areas by reducing the size of the instantaneous data base on which the algorithm operates. This is accomplished by processing data within a "windowed" sub-image, and by stepping the "window" horizontally and vertically through the entire image, adapting the algorithm at each step.

Within each windowed sub-image of 32 lines and 32 pixels per line, a sub-image intensity histogram is formed by tabulating the number of occurrences of each pixel intensity as a function of pixel intensity. Since the existing LAHE computer program operated on video with 6 intensity bits per pixel, the original SAR data were first converted to this format. A sub-image histogram might appear as in Figure 11.

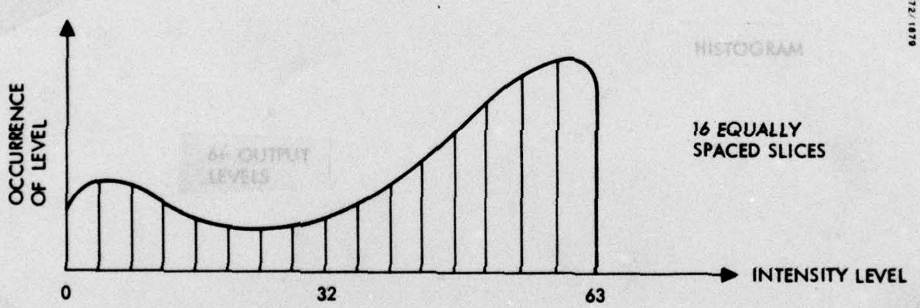


Figure 11. Equal intensity level slicing.

If the sub-image is displayed on a cathode ray tube which generates 16 shades of gray, the result is equivalent to slicing the gray scale axis of the histogram into 16 equally spaced levels. However, if the amount of information conveyed by each gray level is proportional to the number of occurrences of that gray level, then not all shades of gray in the displayed image carry the same amount of information. To maximize the information in each gray level from an entropy point of view, all gray levels should have an equal probability of occurrence. This can be achieved by reassigning the intensity

14

boundaries between displayed gray shades such that all display gray shades occur approximately the same number of times. Thus, if 16 gray levels are displayed, each one should occur about $\frac{32 \times 32}{16}$ or 64 times within each sub-image of 32 lines with 32 pixels per line. This means that the areas under the histogram curve between successive gray level intensity boundaries should all be about equal, as shown in Figure 12.

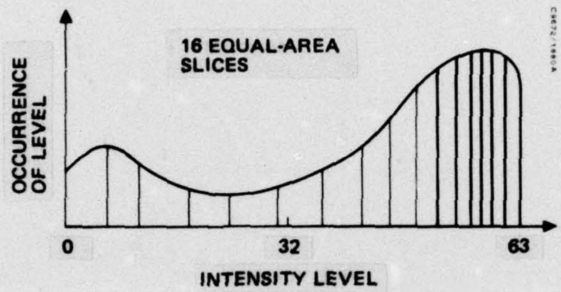


Figure 12. Equal area slicing.

The LAHE process computes the sub-image histogram, finds the equal-area gray shade intensity boundaries, and then assigns a new gray shade to the center element of the 32 by 32 window based on the new gray shade intensity boundaries. The window position is then moved to the right, one element at a time until the current line has been processed, at which time the window moves down to the left edge of the next line and steps across.

The process of forming a histogram around each picture element is quite time consuming. It would be more efficient to compute a histogram and equalize a group of picture elements rather than just one. A trade-off study was performed to determine the optimum number of elements to equalize within each window area. Based upon the histogram in the sliding window of 32 rows by 32 columns, the picture elements in the center N rows and N columns were adjusted, for N = 1, N = 2, N = 4, N = 8, N = 16, and N = 32. It was found that for N = 8, 16, and 32, the enhanced picture has a "boxy" look, i.e., the algorithm adaptation boundaries are visible, while for N = 1, 2, and 4, this effect is absent. Therefore, the conclusion was that the pixels in the center four rows and four columns of each sliding window could be equalized simultaneously without undue picture distortion while producing good computational speed.

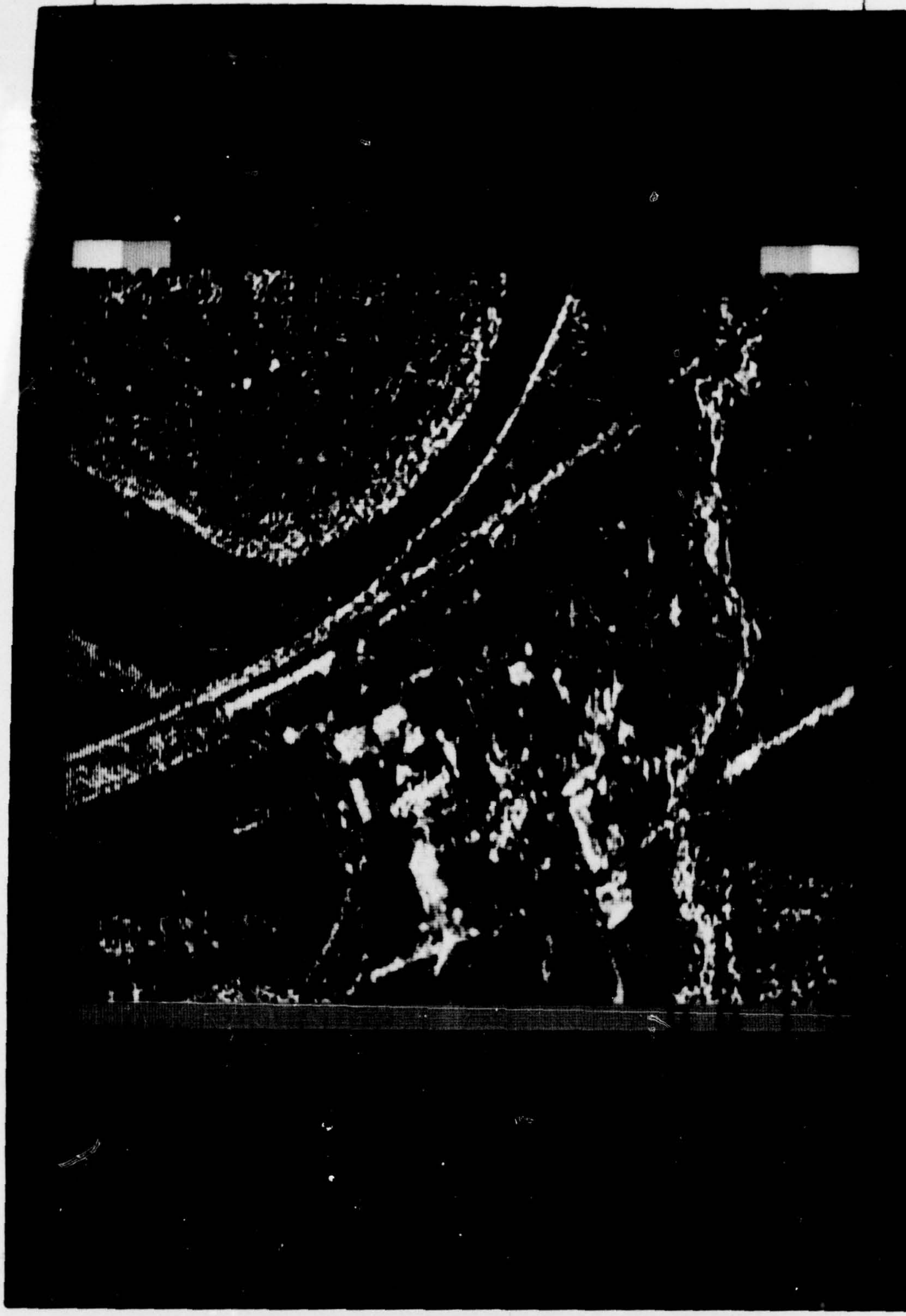
The LAHE technique was therefore performed with $N = 4$ to obtain representative and cost-effective performance. The results are shown in Figure 13. It can be seen that LAHE improves small area contrast. This can be seen in the tree area above and to the right of the bridge scene and at the right edge of the train scene of Figure 13. LAHE is subject to the same criticism directed at the LABGC technique, namely the amplification of clutter. For this reason, neither of these two techniques were considered further. With the addition of target area detection logic to allow the processor to decide when and when not to apply its algorithm, such techniques might be more useful. But the presence of such logic would place the technique in the class based on feature analytic methods, so further discussion will be deferred to Section 4 of this report.

One-Dimensional Haar Stockham Nonlinear Filter

The nonlinear filtering process originally derived by Stockham (1972) is based on a specific mathematical model of what comprises the image intensities. The model considers the energy imaged by a sensor to be proportional to the amount of energy which illuminates the target and to the reflectance characteristics of the target. Thus a pixel at position (x, y) has an intensity $I(x, y) = i_{x,y} \cdot r_{x,y}$, where $i_{x,y}$ is the illumination component and $r_{x,y}$ is the reflectance component. Taking the logarithm of $I(x, y)$, we obtain $D(x, y) = \log I(x, y) = \log i_{x,y} + \log r_{x,y}$. Generally, the illumination component will be diffuse and smooth so that its logarithm will be made up mainly of low frequency components in the transform domain. The reflectance component, however, varies greatly from one target element to the next, and it is this property which provides the image detail. For this reason, the logarithm of the reflectance component will be heavily made up of high frequency values in the transform domain. To enhance the detail due to the reflectance components, the high frequency terms should be emphasized and the low frequency terms attenuated to correspondingly attenuate the overall illuminance.

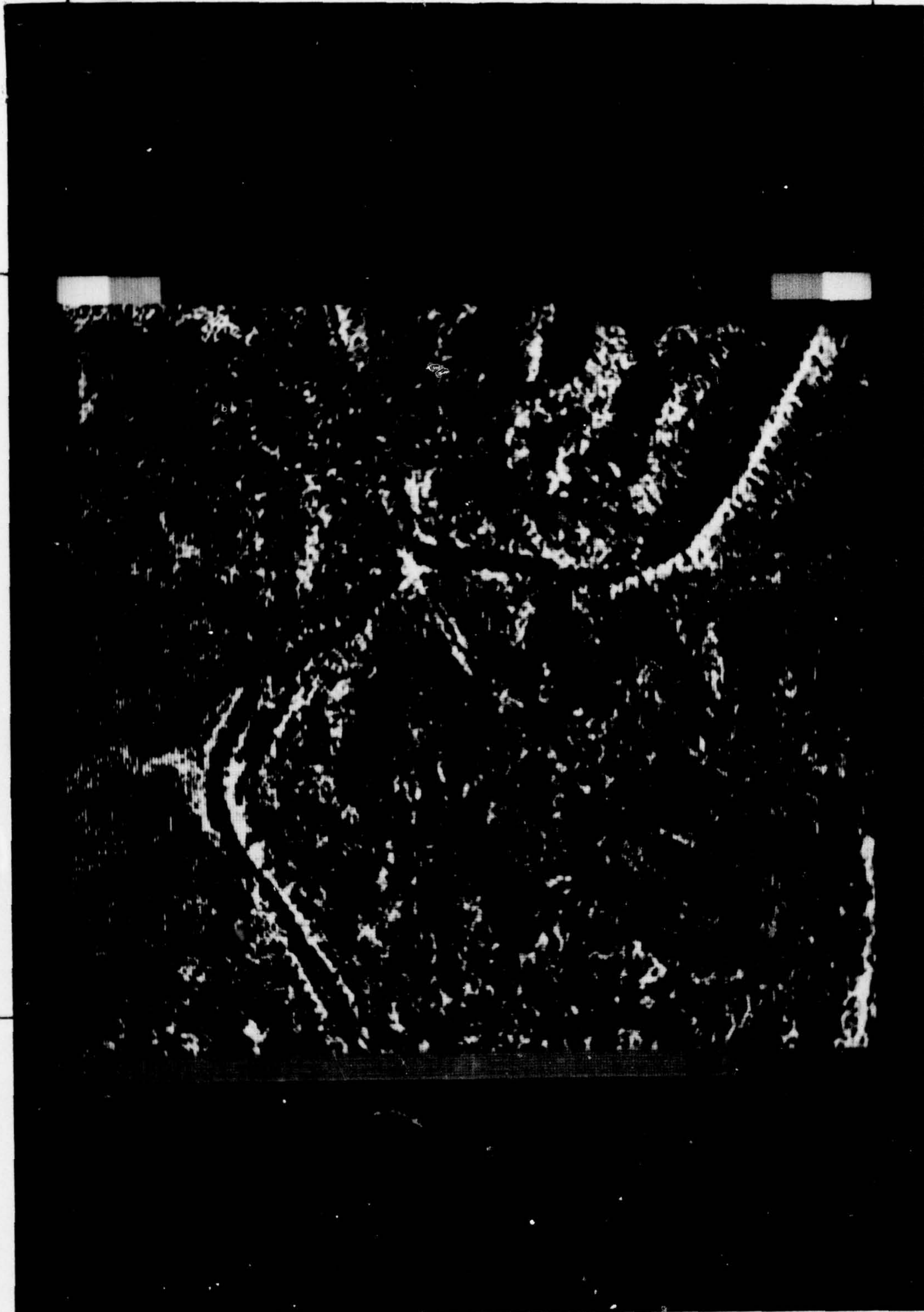
By taking the Fourier transform, for example, of $D(x, y)$, we obtain $F[D(x,y)] = F[\log i_{x,y}] = F[\log r_{x,y}]$, and these two components should inherently be separated in frequency. Thus it is easy to construct $AF[\log i_{x,y}] + EF[\log r_{x,y}] = F[A\log r_{x,y}] = F[E\log r_{x,y}] = F[A\log i_{x,y} + E\log r_{x,y}] = F[D(x,y)]$, and taking the inverse transform

16



F.13d

To 79%



F. 13b

to 79%

1

2

Original

Enhanced

3

4

Original

Enhanced

Figure 13. Local area histogram equalization enhancement.

$$\hat{D}(x, y) = A \log i_{x, y} + E \log r_{x, y}$$

is obtained where A and E are multipliers signifying the attenuation of illuminance and emphasis of reflectance, respectively. If the image is exponentiated, an enhanced image $\hat{I}$ is obtained by

$$\hat{I}(x, y) = \exp(\hat{D}_{x, y}) = i_{x, y}^{A} \cdot r_{x, y}^{E} .$$

Stockham's original work employed a filter in the Fourier domain. Because the Haar transform can be computed with additions and subtractions only, it has greater potential for real-time mechanization than does the Fourier transform, which requires numerous multiplications. Thus, a Haar domain filter equivalent in function to the Fourier filter originally used by Stockham was derived.

The generic one-dimensional Haar non-linear filtering process is described as follows. Let

I = column of pixels in original image (2n x 1 vector)

H = Haar transform operator (2n x 2n matrix)

H^{-1} = inverse Haar transform operator (2n x 2n matrix)

S = Haar domain filter characteristic (2n x 2n matrix)

$\hat{I}$ = column of pixels in enhanced image (2n x 1 vector)

First, a logarithmic conversion is made, followed by a Haar transformation,

$$H \left[\log (I) \right] ,$$

where the notation log (I) denotes a 2n x 1 vector, each of whose elements is the logarithm of the corresponding element of I. The Haar transform of the logarithm of the image is then multiplied by filter coefficients, inverse transformed and exponentiated, i.e.,

$$\hat{I} = \exp \left[H^{-1} \cdot S \cdot H \cdot \log (I) \right] ,$$

where the notation exp [A] denotes a 2n x 1 vector, each of whose elements is the exponential of the corresponding element of A. This process is applied to each column of the original image until the entire image has been processed.

It was decided to first examine these filtering techniques in one dimension only, because to perform two-dimensional filtering requires that the entire process described in the preceding paragraph be repeated on the rows of the one-dimensionally filtered image. Filtering in the second dimension requires image transposition. Most SAR images of interest have enough resolution and coverage that the image data must reside in secondary storage, such as on a disk. In this case, image transposition is a time consuming process, requiring many input/output operations from and to the disk. Because of this, the processing time for two-dimensional filtering is much more than a factor of two greater than the processing time for one-dimensional filtering.

The transfer function magnitude of Stockham's filter is shown in Figure 14. It basically has a high pass characteristic, reflecting the intent to attenuate low spatial frequencies and emphasize higher spatial frequencies.

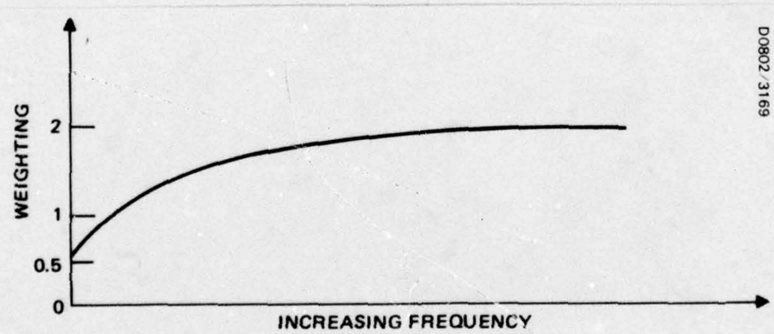


Figure 14. Stockham Fourier filter transfer function magnitude.

Computations were performed to obtain a matrix corresponding to the transfer function shown in Figure 14. For computational simplicity, the off-diagonal terms in the resulting matrix were set to zero, yielding the approximate Haar domain equivalent filter characteristic shown in Figure 15.

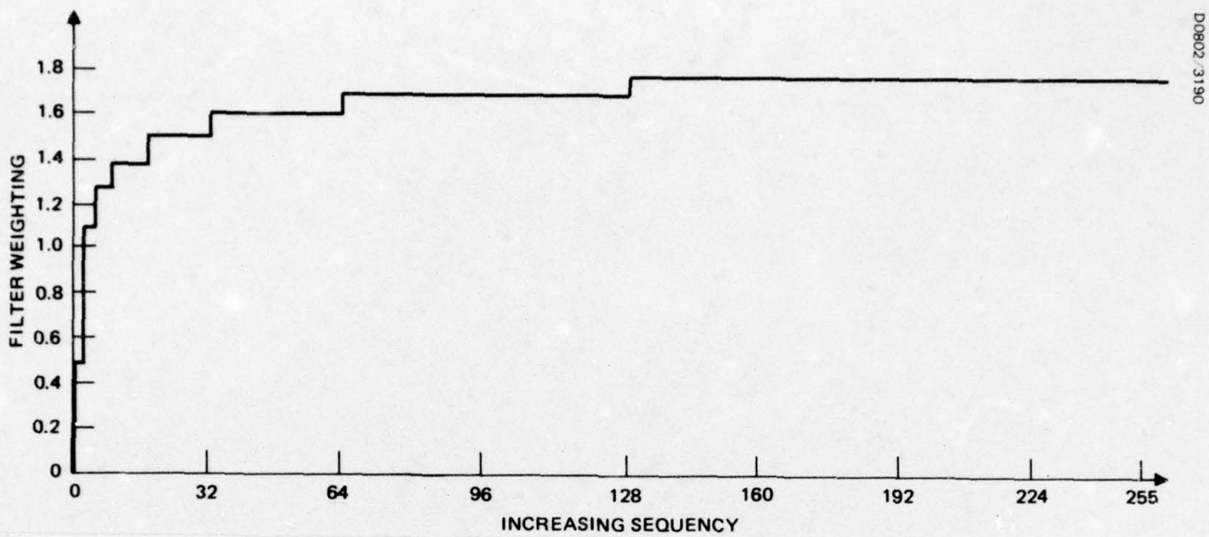


Figure 15. Haar domain filter derived from the Stockham Fourier filter.

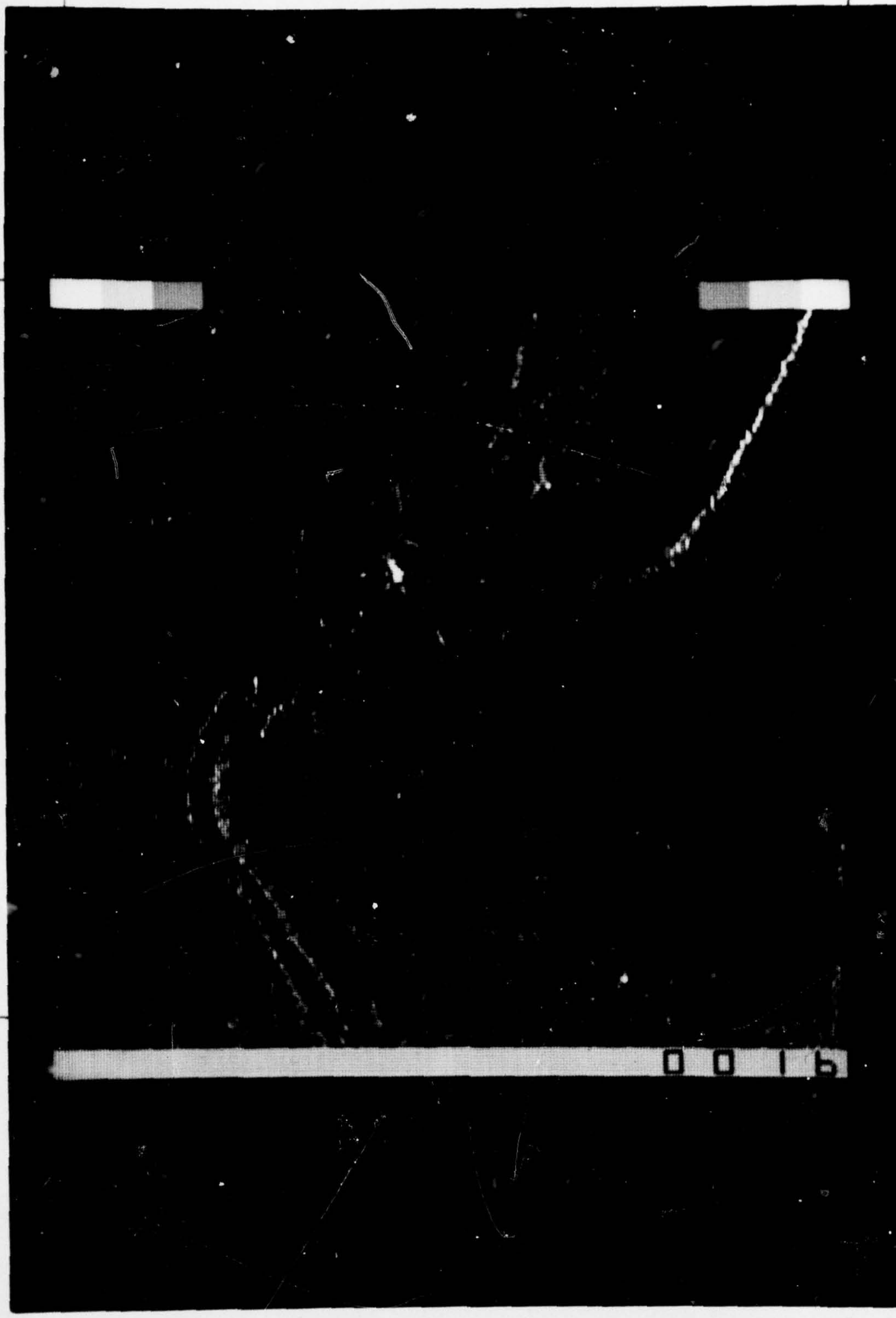
The computer program which performed Stockham non-linear filtering required the input image data to be in a 6-bit per pixel format, so the conversion from floating point to 6-bit fixed point representation was first made. The Haar domain filter was then applied. The results are shown in Figure 16. F 16

Image contrast has been improved, as evidenced by the convoy at the lower right of the bridge scene and the bright point returns at the upper left of the train scene. Also, the clutter emphasis effects present in LABGC and LAHE are much diminished in Figure 16. The amount of operator performance improvement, however, is not expected to be of any practical significance. There is an overall gain in image enhancement with the Stockham filter, but not enough to produce improved operator performance.

One-Dimensional Haar Non-Linear Crispening Filter

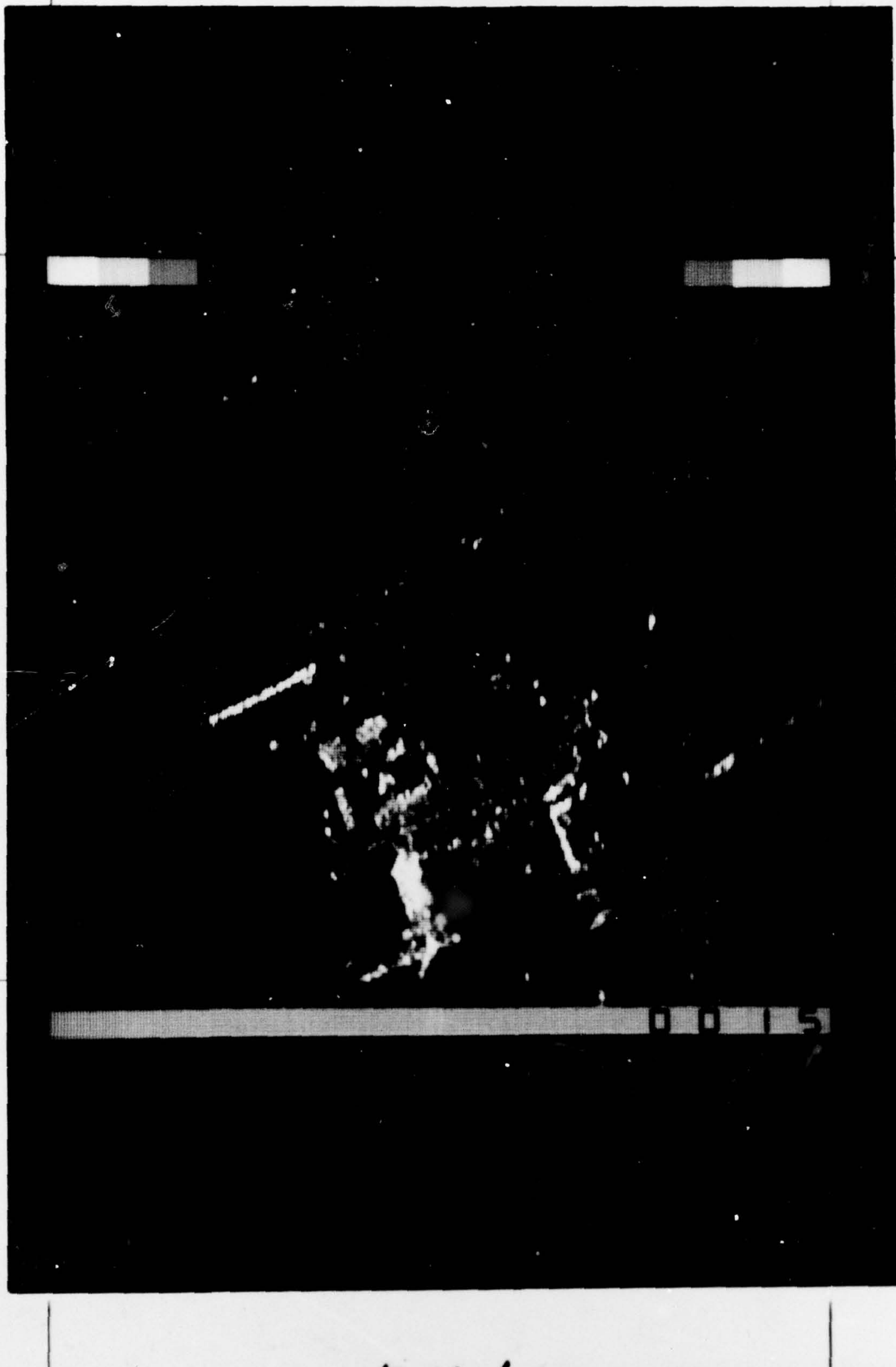
Since edge information is important in target location tasks, it could prove desirable to emphasize edges in SAR images. This can be done in one dimension by a second derivative operator in the form $I - kD_x^2$, where k is a gain constant and I and D_x^2 are the identity and one-dimensional spatial second derivative operations, respectively. Figure 17 illustrates the significance of subtracting a weighted second derivative from a function. Figure 17a represents the spatial profile of image intensities found in transitioning the boundary of a white and a black area. Figure 17c shows that F 17

20



F. 16b

to 79%



F.16d

to 79%

1

2

Original

Enhanced

3

4

Original

Enhanced

Figure 16. One-dimensional Haar Stockham non-linear filter enhancement.

21

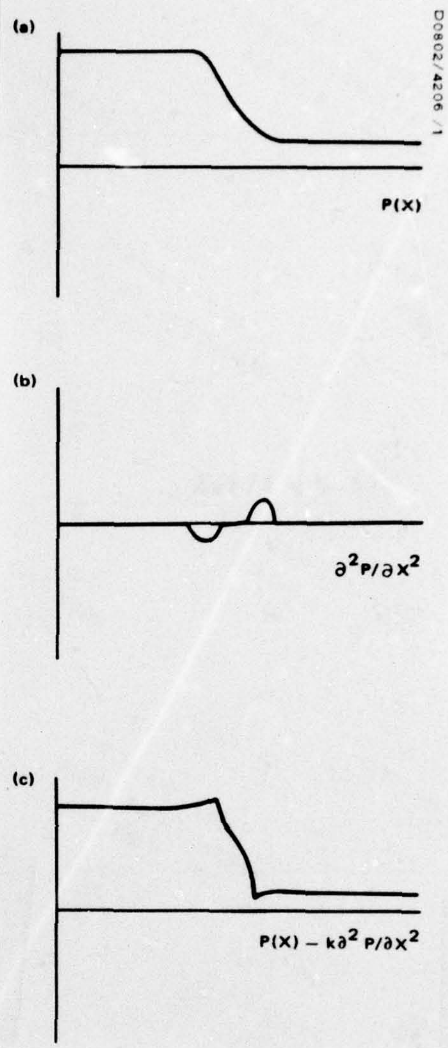


Figure 17. Second derivative edge crispening.

subtracting some of the second derivative from the original function increases the magnitude of the transition from peak white to peak black and decreases the spatial separation of the local image intensity extrema. These two effects contribute to the perception of a sharper edge. Such differentiation operations in the spatial domain are highly susceptible to image noise. An operator was therefore derived to perform the differentiation in the Haar transform domain. The reasoning was that since the Haar transform coefficients are computed from sums and differences of the image data, some of the noise would be averaged out in the transform coefficient computations. The approach taken was the same as that used to derive the Haar equivalent of the Stockham Fourier filter. It can be shown that the Fourier domain filter which performs a second derivative operation is a diagonal matrix whose elements increase in magnitude proportional to the square of the corresponding spatial frequency. Such a characteristic is shown in Figure 18.

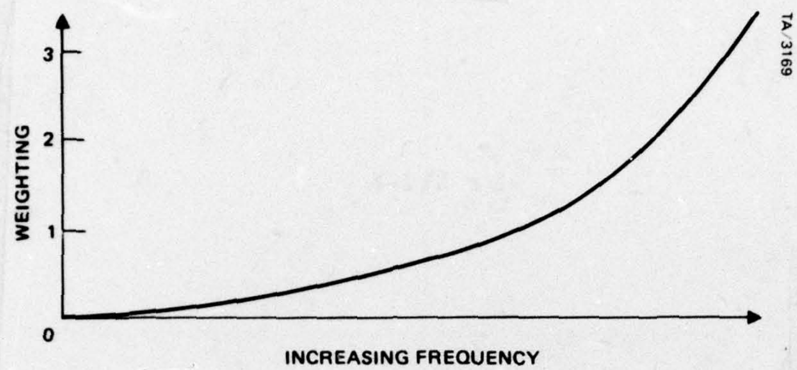


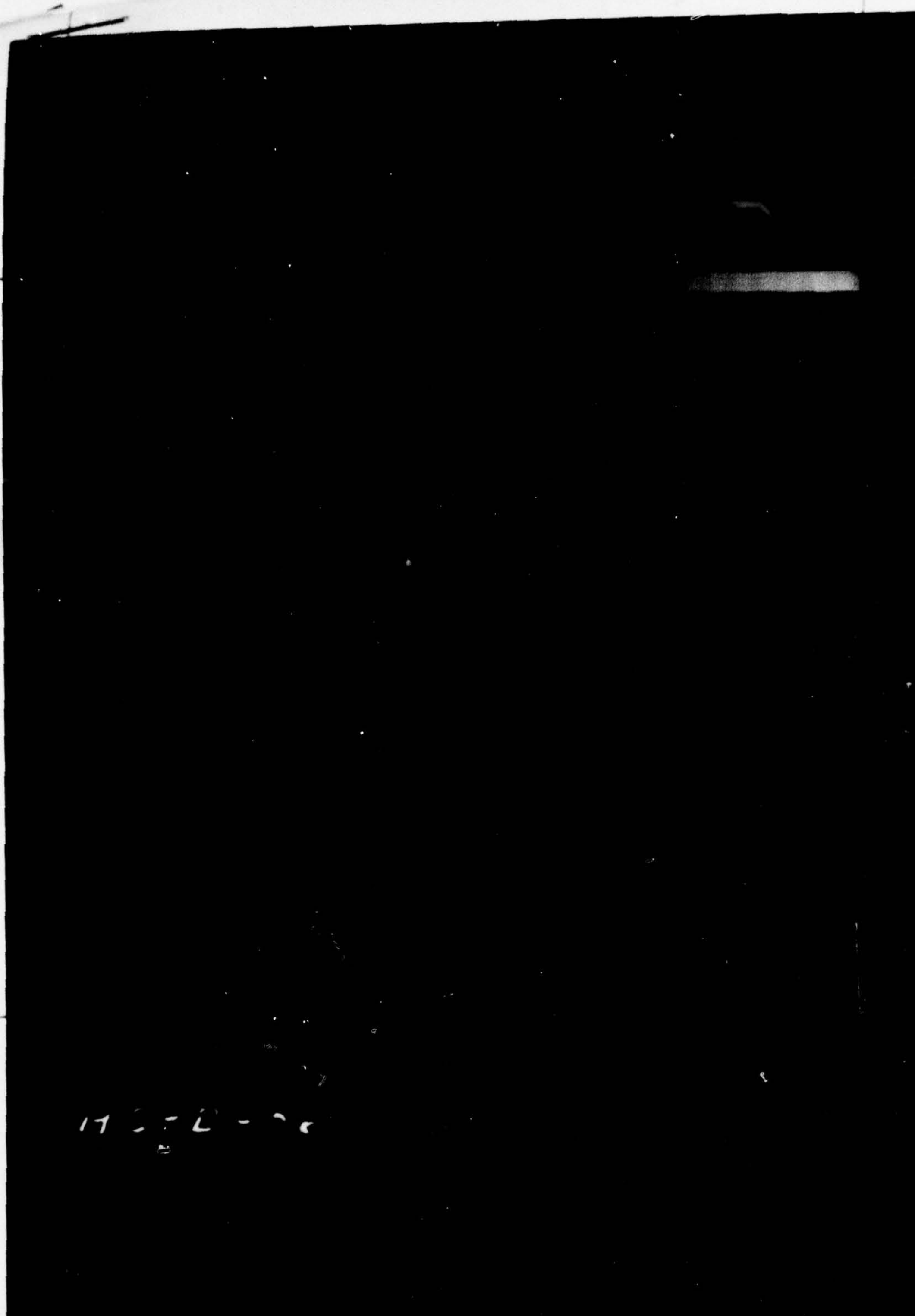
Figure 18. Fourier second derivative filter characteristic.

By the use of the change of basis matrices, an equivalent Haar domain filter was obtained.

As with the Stockham non-linear filtering computer program, the non-linear crispening program required the input image to be in 6-bit per pixel format. The necessary conversion was performed prior to filtering. The images then underwent the logarithm operation, were filtered in one dimension with the non-linear crispening filter, and then exponentiated. Figure 19 shows the processed images obtained using a gain constant of one. As can be seen from the black dots at the edges of the highest brightness objects, particularly in the train scene, the steepness of edge transitions has been increased. Although not obvious in Figure 19, an investigation of the sensitivity of this technique to image noise (Ketcham, Lowe, and Weber, 1976) indicated that Haar transform edge crispening is more noise sensitive than LABGC or LAHE, particularly with radar imagery. For this reason, and because the increase in target conspicuity with this process is marginal, it was decided that further modification was necessary to make such a technique viable. These modifications are described later.

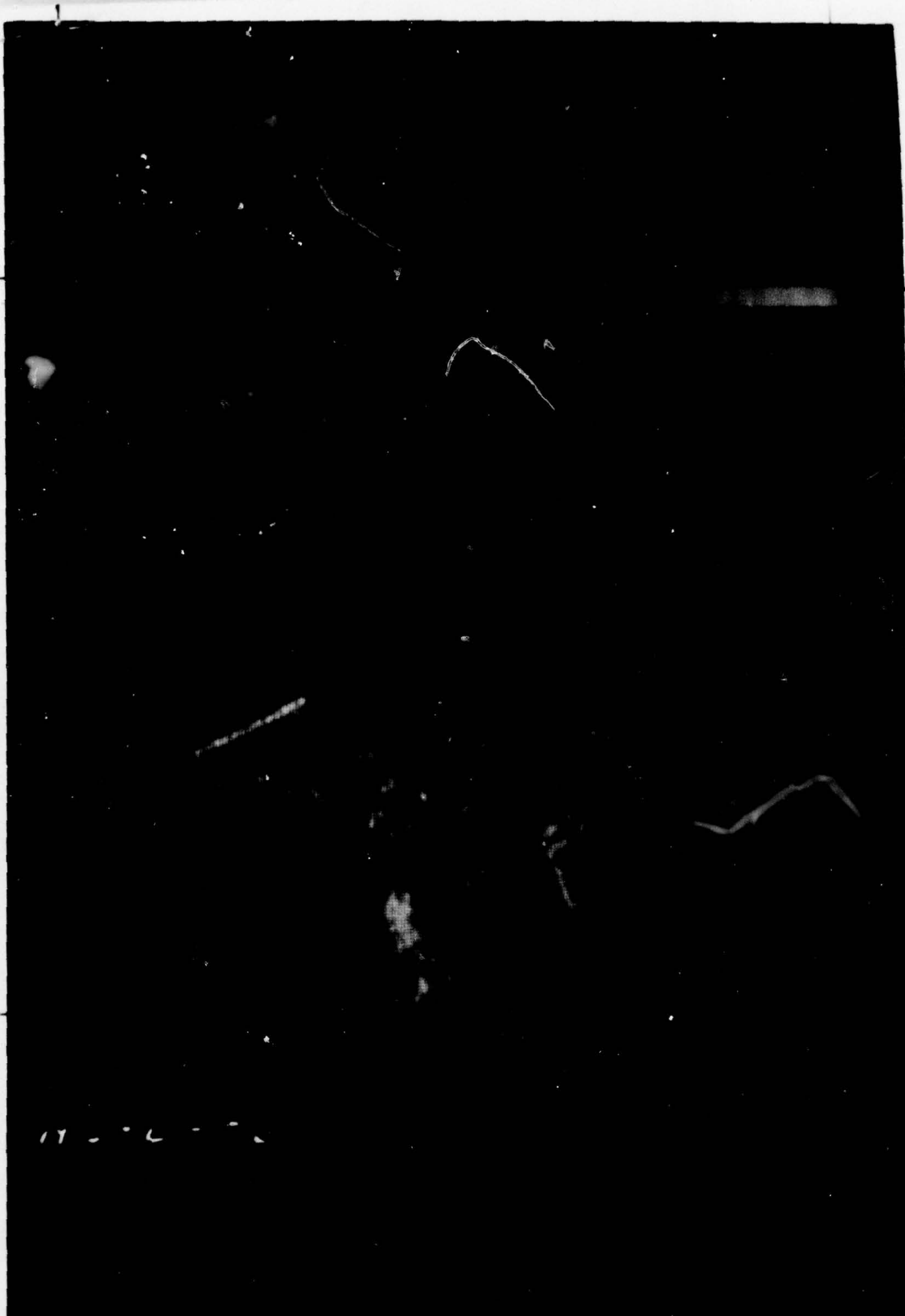
One-Dimensional Haar Linear High Pass Filter

Recall from the discussion of the Haar Stockham non-linear filter that the image data stored on the disk in floating point format was converted



F. 19b

407920



F. 19d

to 7990

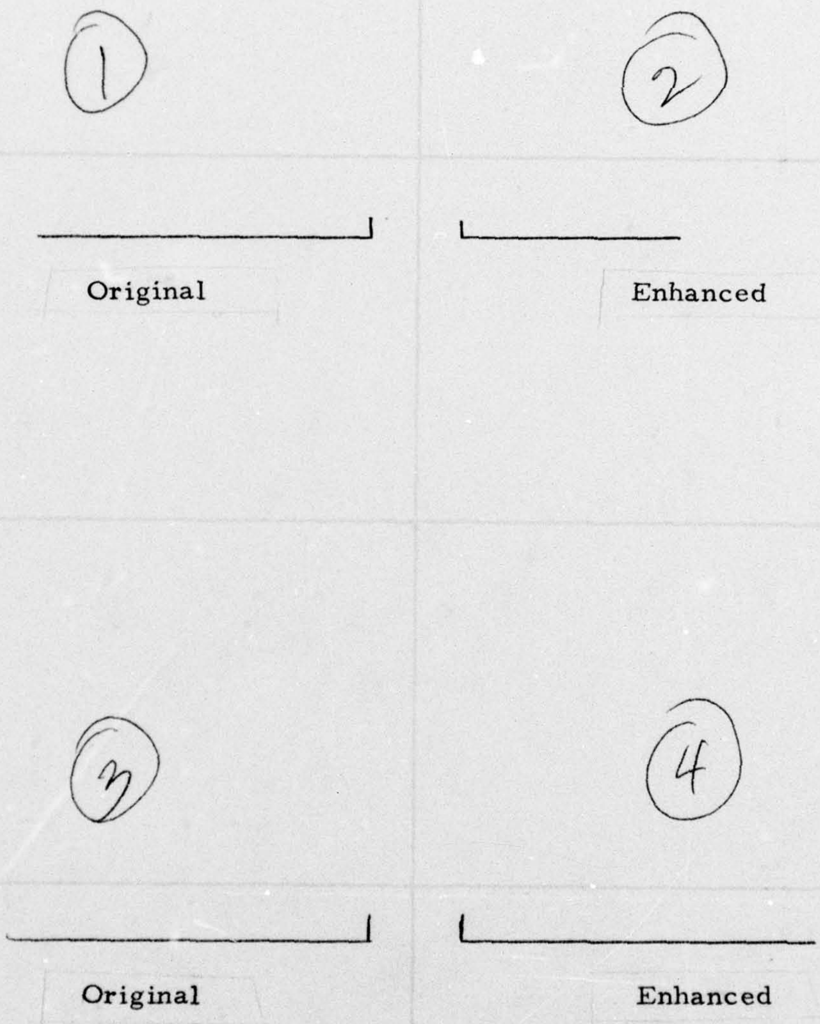


Figure 19. One-dimensional Haar non-linear crispening enhancement.

to 6-bit per pixel format — a logarithmic operation — prior to filtering. The first step of the non-linear filtering process is to perform another logarithm operation; hence, two logarithm operations were performed on the floating point image data, as shown in Figure 20. If the 8-bit fixed point data are considered to be the original image, then the additional logarithm step of the non-linear filtering process is appropriate to perform Stockham non-linear filtering on the 8-bit data. However, if the floating point data on the disk are considered to be the original image, then the second logarithm step is not necessary to achieve Stockham non-linear filtering of that original data.

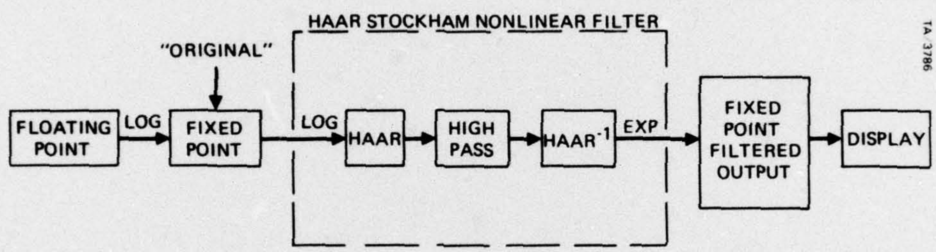


Figure 20. Block diagram of the Haar Stockham non-linear filtering process for fixed point original data.

If the floating point data are considered to be the original image, then after the logarithmic conversion to fixed point and the spatial frequency filtering steps, the fixed point output should be exponentiated to floating point format for comparison with the original image. Since floating point data can not be displayed, the logarithm would have to be taken to convert back to fixed point for display. Thus, the final exponentiation and logarithm steps are both omitted, since they would cancel each other. Figure 21a illustrates the theoretically necessary steps, while Figure 21b indicates what was actually performed. It can be seen that the two sequences are equivalent. F21

The high pass filter characteristic shown in Figure 15 was used in the linear filtering investigation. Only the logarithm and exponentiation operations shown in Figure 20 for the non-linear filtering sequence were omitted; otherwise, all filtering parameters were the same as those used in the examination of Haar Stockham non-linear filtering. The results are shown in Figure 22. F22

25

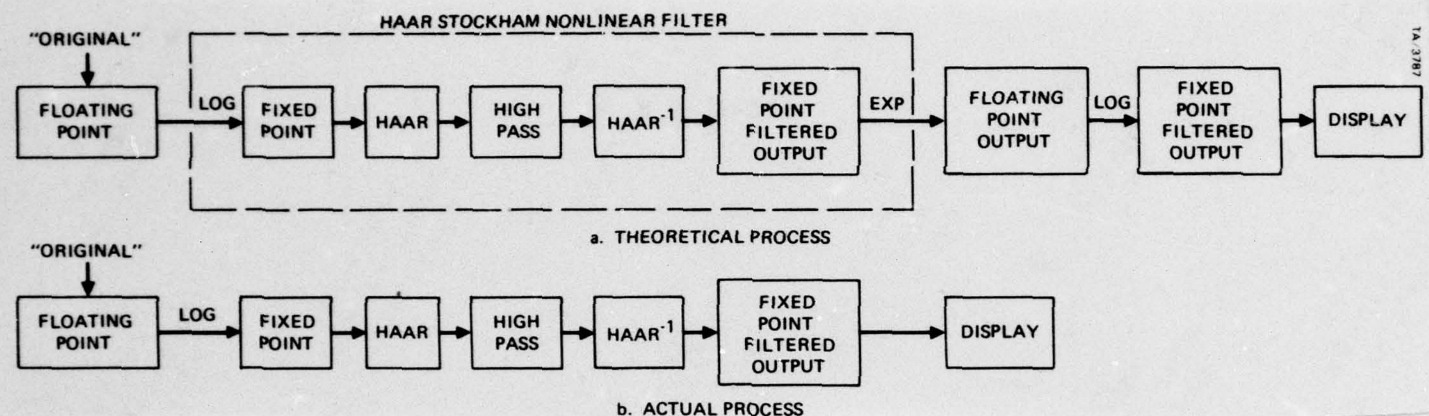
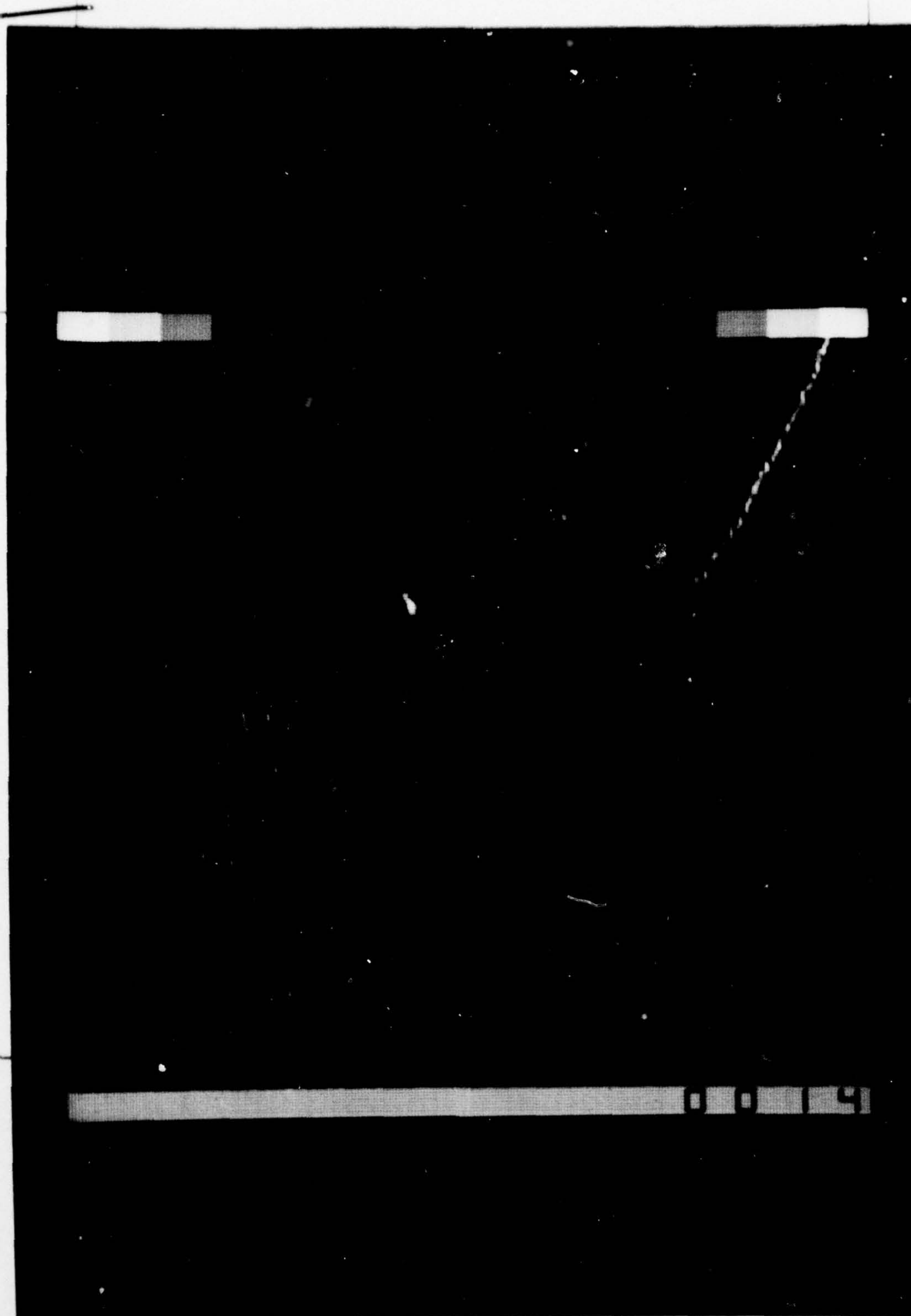
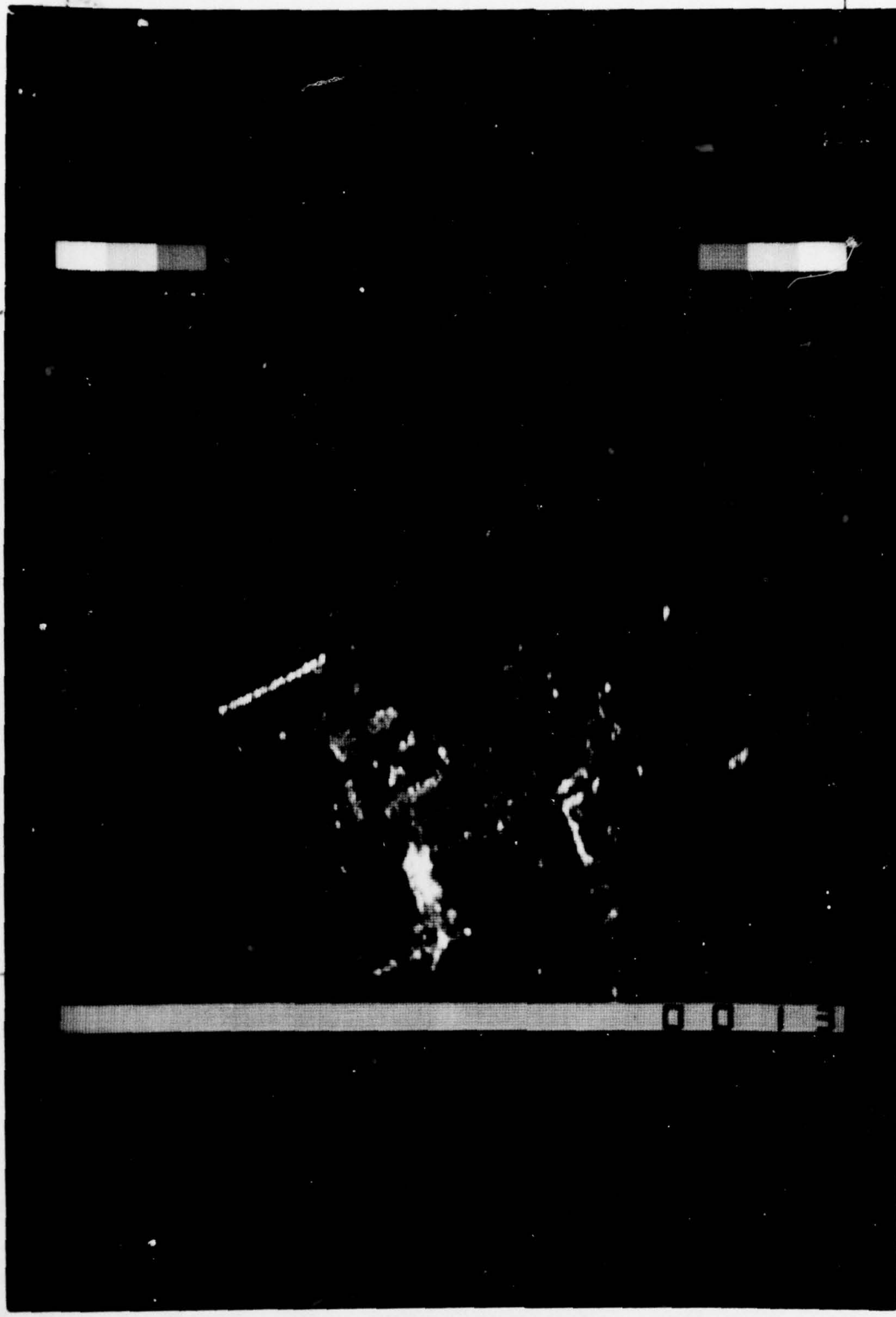


Figure 21. Block diagram of theoretical and actual Haar Stockham non-linear filtering process for floating point original data.



F.22b

to 79%

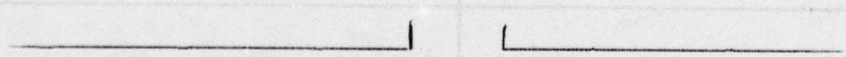


F.22d

to79%o

1 2

Original Enhanced



Original Enhanced

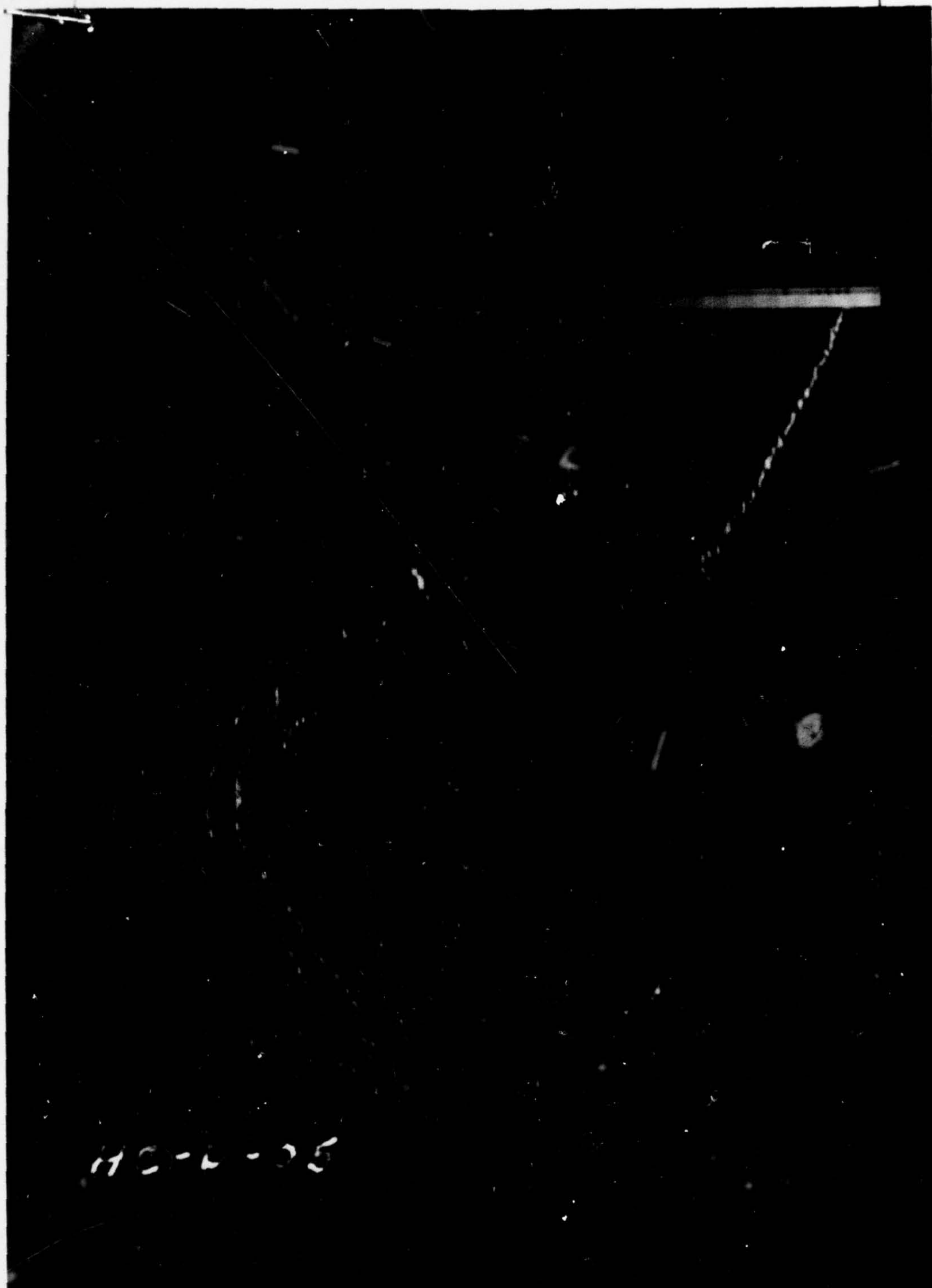
Figure 22. One-dimensional Haar linear high pass filter enhancement.

An increase in contrast and edge sharpness due to the relative emphasis of high spatial frequencies over low spatial frequencies of the high pass filter is noticeable in the enhanced images. Comparison with the non-linearly filtered images in Figure 16 indicates little practical difference; although, the non-linear process may provide somewhat better enhancement in the low brightness areas of the image. Such marginal differences do not seem to be worth the additional cost of non-linear processing. Linear high pass filtering does not have the clutter emphasis problems of LABGC or LAHE and is about the same as Stockham non-linear filtering in this regard.

One-Dimensional Haar Linear Crispening Filter

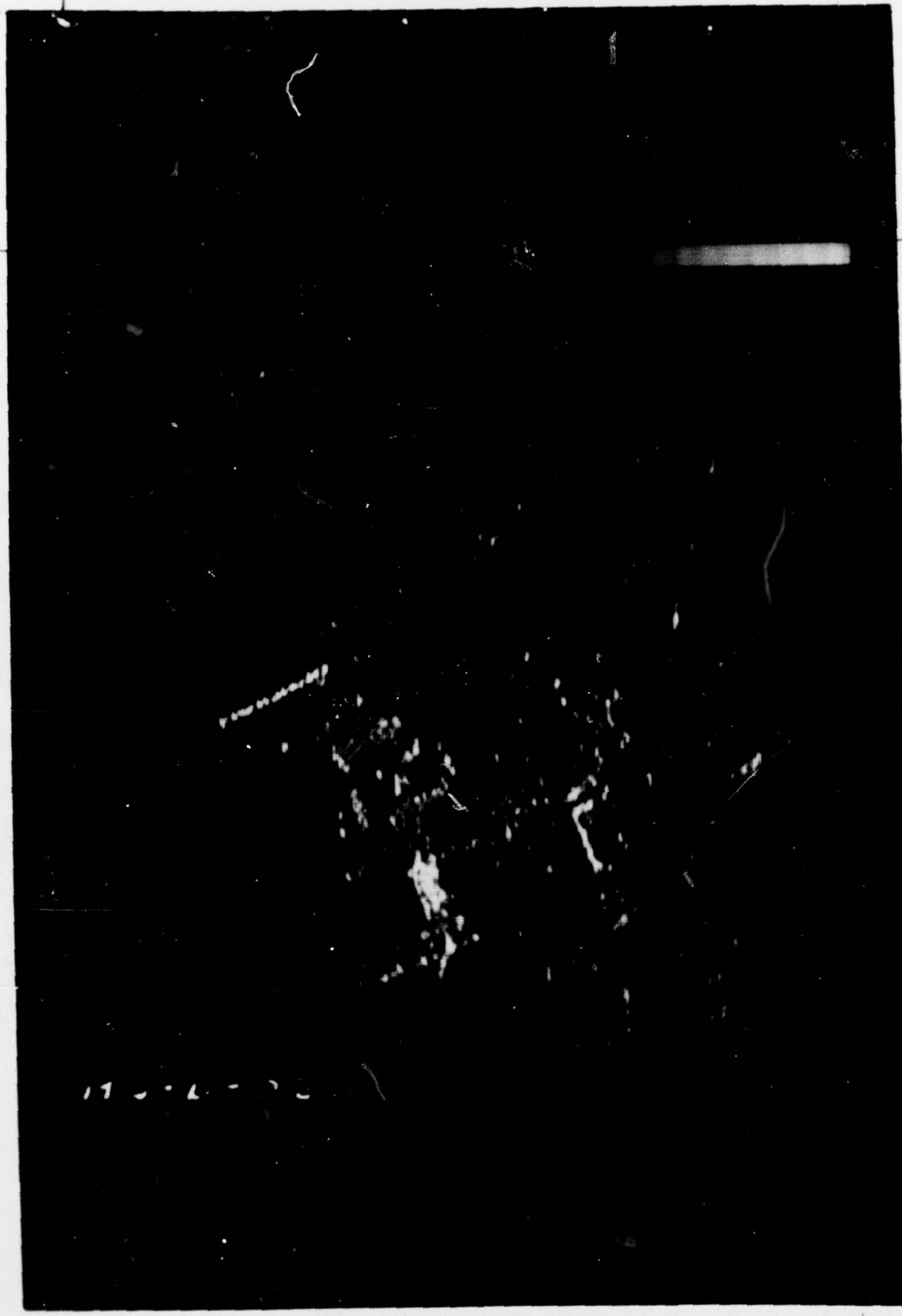
The rationale which led to the investigation of Haar linear high pass filtering applies as well to the process of Haar non-linear crispening. If the words "Stockham non-linear filter" are replaced by "non-linear crispening filter" and the words "high pass" are replaced by "crispening" throughout the discussion of one-dimensional Haar linear filtering, one obtains the rationale for Haar linear crispening.

As with high pass filtering, the same crispening filter characteristic was used for both linear and non-linear processing; namely, the filter characteristic shown in Figure 18. A gain value of $k = 0.5$, however, was employed in the linear processing. The linearly processed images are shown in Figure 23. The linearly crispened images appear to have sharper edges than do the originals. There is, however, a tendency toward graininess which would become more pronounced as the signal-to-noise ratio of the original images was decreased. Relative to the non-linearly crispened images in Figure 19, the linearly crispened images appear to have better contrast and edge sharpness. This is probably due to the edge transition compression effect of the logarithm operation prior to crispening in the non-linear process. The subsequent exponentiation is then unable to make up for the loss in contrast produced by smaller edge transitions due to taking the logarithm. Finally, in comparison to the linearly high pass filtered images in Figure 22, the linearly crispened images have about the same contrast, somewhat better sharpness, and somewhat more granularity. The greater granularity implies poorer noise sensitivity properties.



F. 23b

to79%



F. 23d

+079%

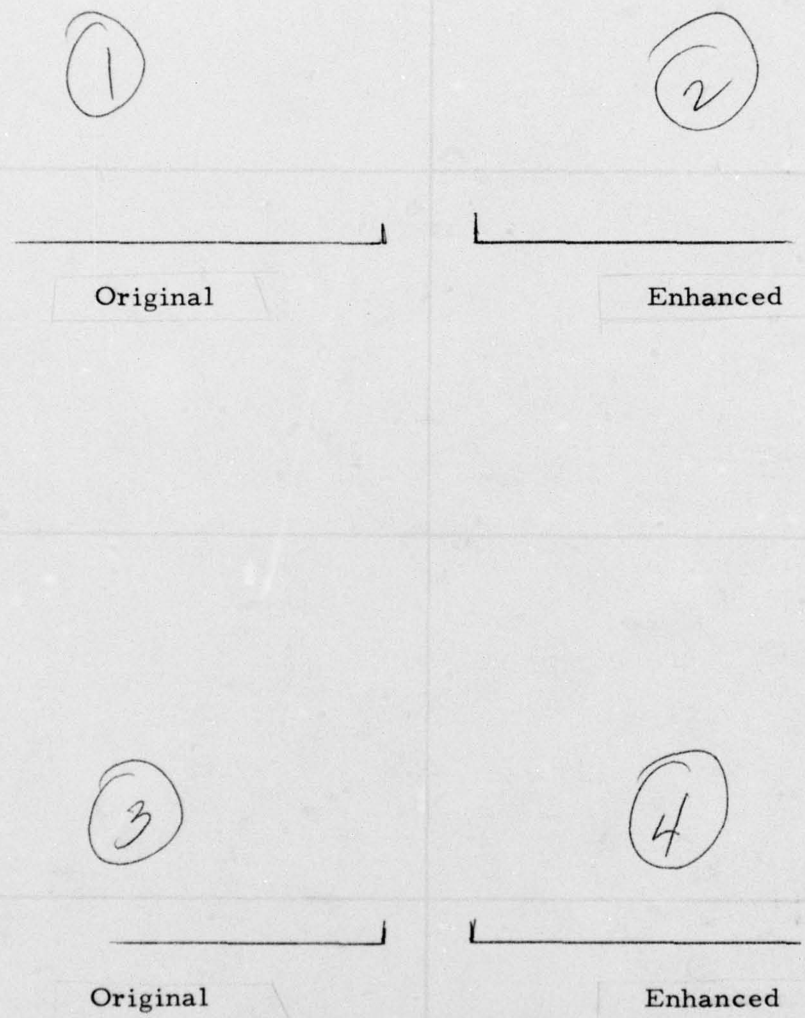


Figure 23. One-dimensional Haar linear crispening enhancement.

29

One-Dimensional Fourier Modified Crispening Filter

The four preceding image enhancement techniques were implemented in the Haar transform domain because the Haar transform required operations which are fewer in number and simpler in nature than the Fourier transform. A consideration that tends to offset the computational advantage of the Haar transform is that some approximation is usually necessary in the Haar domain to maintain the computational simplicity of the Haar domain filter. This is because the Haar filter which is equivalent to a desired Fourier domain filter characteristic frequently has a large number of non-zero off-diagonal terms, even when the Fourier filter has none. These non-zero off-diagonal terms are often ignored to simplify the filter mechanization, but the resulting filter is not an exact equivalent of the desired characteristic. In the interest of avoiding such approximation problems and any potential loss of image enhancement performance that might be associated with them, the remaining five image enhancement techniques were all implemented in the Fourier domain.

As found in a noise sensitivity study of LABGC, LAHE, and Haar domain linear and non-linear crispening (Ketcham, Lowe, and Weber, 1976), the crispening process had the highest susceptibility to image noise. This is not surprising in view of the increase in crispening filter gain in proportion to the square of spatial frequency. Reasoning that the effect of wide-band image noise is more noticeable at high spatial frequencies (scintillation noise, radar glints, etc.) than at low or middle spatial frequencies, it was conjectured that a better crispening filter might be obtained by rolling off the filter gain above some cutoff spatial frequency. The cutoff frequency should be chosen so that most of the target spatial frequencies still are crispened. Such a filter characteristic is a modified crispening filter.

A generic filter characteristic for the modified crispening filter is shown in Figure 24. The parameter FREQ is the filter cutoff frequency in cycles per image dimension; BAND is the frequency interval over which the filter gain rolls off beyond cutoff in cycles per image dimension; LEVEL is the filter gain for frequencies greater than FREQ + BAND, and GAIN is a scale factor which multiplies the portion of the filter characteristic corresponding to frequencies below FREQ + BAND. In addition, a parameter WFACT was a weighting factor which determined how much of the modified second derivative information was subtracted from the original image data. All these

Fig 24

30

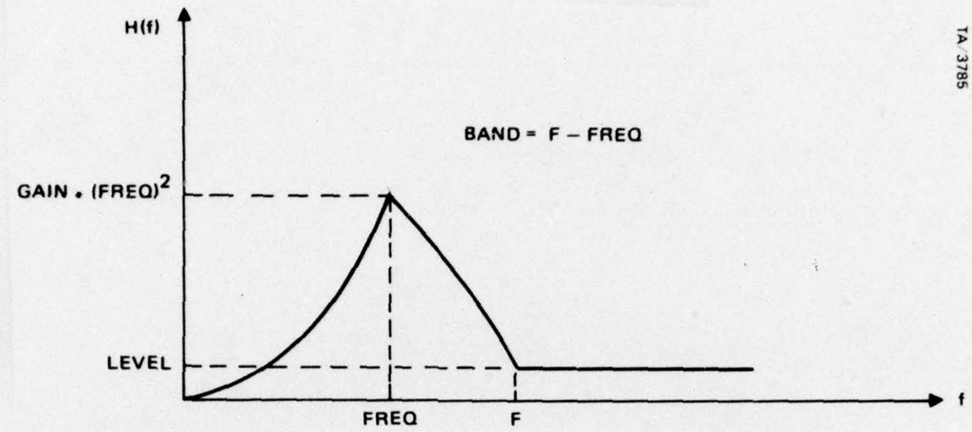


Figure 24. Generic modified crispening filter characteristic.

parameters were independently variable through software. Some optimization of the modified crispening filter with respect to these parameters was performed. The performance of the filter was found not to be sensitive to parameter variations of less than about half an order of magnitude, so that once the correct order of magnitude of a parmeter value was discovered, further identification of "optimum" values within a factor of about three was unnecessary.

The modified crispening filter program, and all the subsequent enhancement technique programs employing filters in the Fourier domain, were written to accept floating point data as stored on the disk as input. The original data were Fourier transformed, multiplied by the modified crispening filter coefficients, and inverse Fourier transformed. This floating point filter output was multiplied by WFACT and subtracted from the original floating point data. The resulting image data were then logarithmically transformed to fixed point mode for display. With the exception of the logarithmic conversion required for display, the Fourier modified crispening filter is an entirely linear operation. Figure 25 shows representative results of the modified crispening filter process with parameter values: GAIN = 0.00192, FREQ = 20, BAND = 108, LEVEL = 0.25, and WFACT = 0.25. F25

Comparing the originals with the enhanced images in Figure 25, it is apparent that the enhanced images are somewhat smoother because of the high frequency rolloff. Because of the relative low frequency de-emphasis, the contrast in the enhanced images seems to be slightly better than in the

31

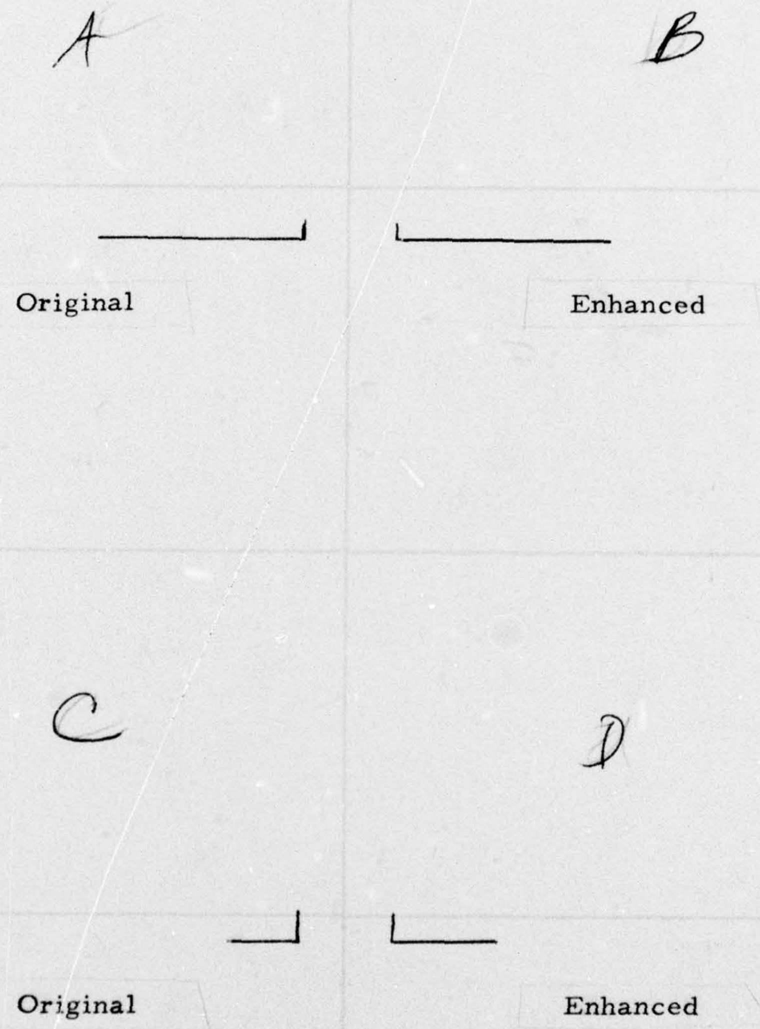


Figure 25. One-dimensional Fourier modified crispening enhancement.

originals. The edge sharpness is about the same as Haar non-linear crispening; although, the modified crispening filter shows better contrast. Comparing the Haar crispening filter (Figure 21) with the modified crispening filter (Figure 25), it is apparent that the high frequency rolloff is effective. The edges are not as sharp, and there is much less image granularity. The objective of obtaining better noise suppression in a crispening type filter was achieved. However, the image enhancement provided by such a filter was not large, and further evaluation of this technique does not appear to be merited.

One-Dimensional Fourier Multilobe Filter

The Fourier multilobe filter was a further attempt to obtain edge enhancement with the additional property that the filter is "tuned" to a target size of interest. The rationale for the development was as follows. A target whose intensity profile in a given spatial dimension is a rectangular pulse of extent n pixels will have a spatial frequency amplitude spectrum (Fourier transform) of the form

$$\frac{\sin \frac{\pi f n}{N}}{\frac{\pi f n}{N}}$$

if there are a total of N pixels along the given spatial dimension. If the absolute value of this spectrum is plotted as a function of f, it is seen that zeros occur at values of f equal to $\frac{kN}{n}$, where $k = 1, 2, 3, \ldots$. Local maxima of the spectrum magnitude occur at $f = 0$ and $\left(\frac{2k+1}{2}\right)\frac{N}{n}$, where $k = 1, 2, 3, \ldots$ ..., taking on values of 1 and $\frac{2}{(2k+1)\pi}$, respectively. The product of the location and the magnitude of the local maxima beyond the first zero of the amplitude spectrum is a constant, i.e. $\left[\left(\frac{2k+1}{2}\right)\frac{N}{n}\right]\left[\frac{2}{(2k+1)\pi}\right] = \frac{N}{n\pi}$, a constant independent of k. This is in effect a $\frac{1}{f}$ rolloff of spatial frequencies.

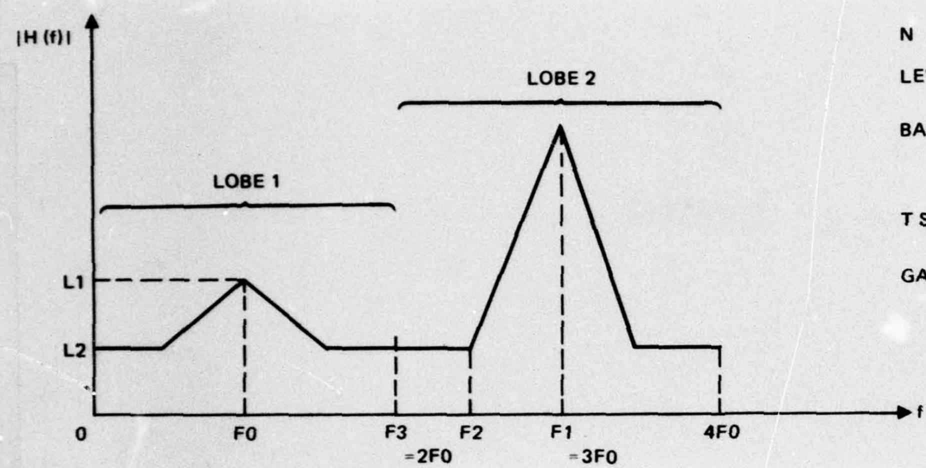
The multilobe filter is intended to achieve edge crispening by combatting the $\frac{1}{f}$ rolloff. It does so by applying a filter characteristic whose local maxima are positioned coincident with those of the amplitude spectrum of a given target of interest beyond the first zero and whose magnitude at these local maxima increases linearly with spatial frequency. Between local maxima, the multilobe filter gain decreases linearly to some low residual level.

Figure 26 shows a generic filter characteristic for the multilobe filter. The number of lobes N is just half of TSIZE, which is the spatial extent in pixels of the target of interest. Thus the filter characteristic shown in Figure 26 corresponds to a target size of 4 pixels, because there are two lobes. The target extent TSIZE also determines the spatial frequency location FO of the first local maximum of the multilobe filter, so if TSIZE = 4, then FO = 32 cycles per image dimension. The parameter BAND specifies the width of the lobes at their base as a fraction of 2FO, which is the maximum possible value; LEVEL is the ratio of the residual relative filter gain to the amplitude of the local maximum of the relative filter gain on the first lobe. The parameter GAIN is a multiplicative amplitude scale factor for the entire filter characteristic. As with the modified crispening filter, TSIZE, GAIN, BAND, and LEVEL were independently variable.

Effort was expended to optimize the parameters GAIN, BAND, and LEVEL. It was found that BAND = 1 consistently gave the best performance, and LEVEL = 0.5 was a good value. The optimum value of GAIN could not be identified quite as accurately, but values on the order of 10^{-2} gave the best results.

To apply the multilobe filter, the original floating point data were Fourier transformed, multiplied by the multilobe filter coefficients, and inverse Fourier transformed. The floating point filter output was then logarithmically transformed to fixed point mode for display. Figure 27 shows representative multilobe filter results, using parameter values GAIN = 0.02, TSIZE = 2 pixels, BAND = 1, and LEVEL = 0.5.

The filtering direction for the images in Figure 27 was horizontal. Since the filter was "tuned" to objects of extent 2 pixels in the horizontal direction, the road is visible in the upper half of the image where it runs transverse to the filtering direction. In the lower half of the image, however, the road is nearly indistinguishable, because it runs more nearly parallel to the filtering direction and has a greater spatial extent in the filtering direction than that size to which the filter is "tuned". It appears that the filter is selectively enhancing those objects of the size to which it is "tuned". The difficulty is that the filter tuning is too selective, i.e., other objects of different sizes, which are important for orientation, context, and similar information, are being tuned out. Due to the resulting



TA/3784

N = NUMBER OF LOBES

$$\text{LEVEL} = \frac{L2}{L1}, \quad 0 < \text{LEVEL} < 1$$

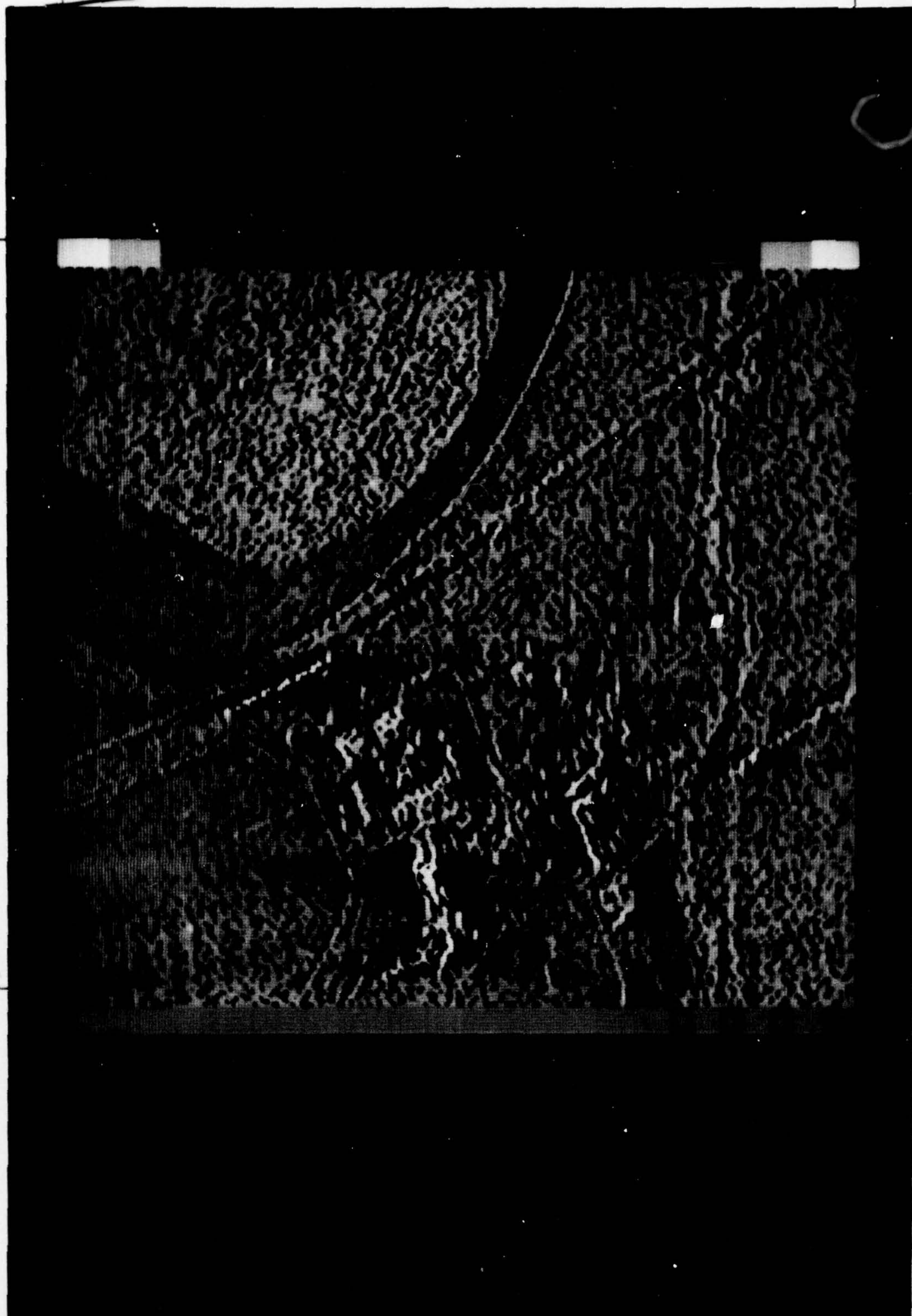
$$\text{BAND} = \frac{F1 - F2}{F1 - F3} = \frac{3F0 - F2}{F0}, \quad 0 < \text{BAND} \leq 1$$

$$\text{T SIZE} = \frac{128}{F0} = 2N$$

GAIN = AMPLITUDE SCALE FACTOR

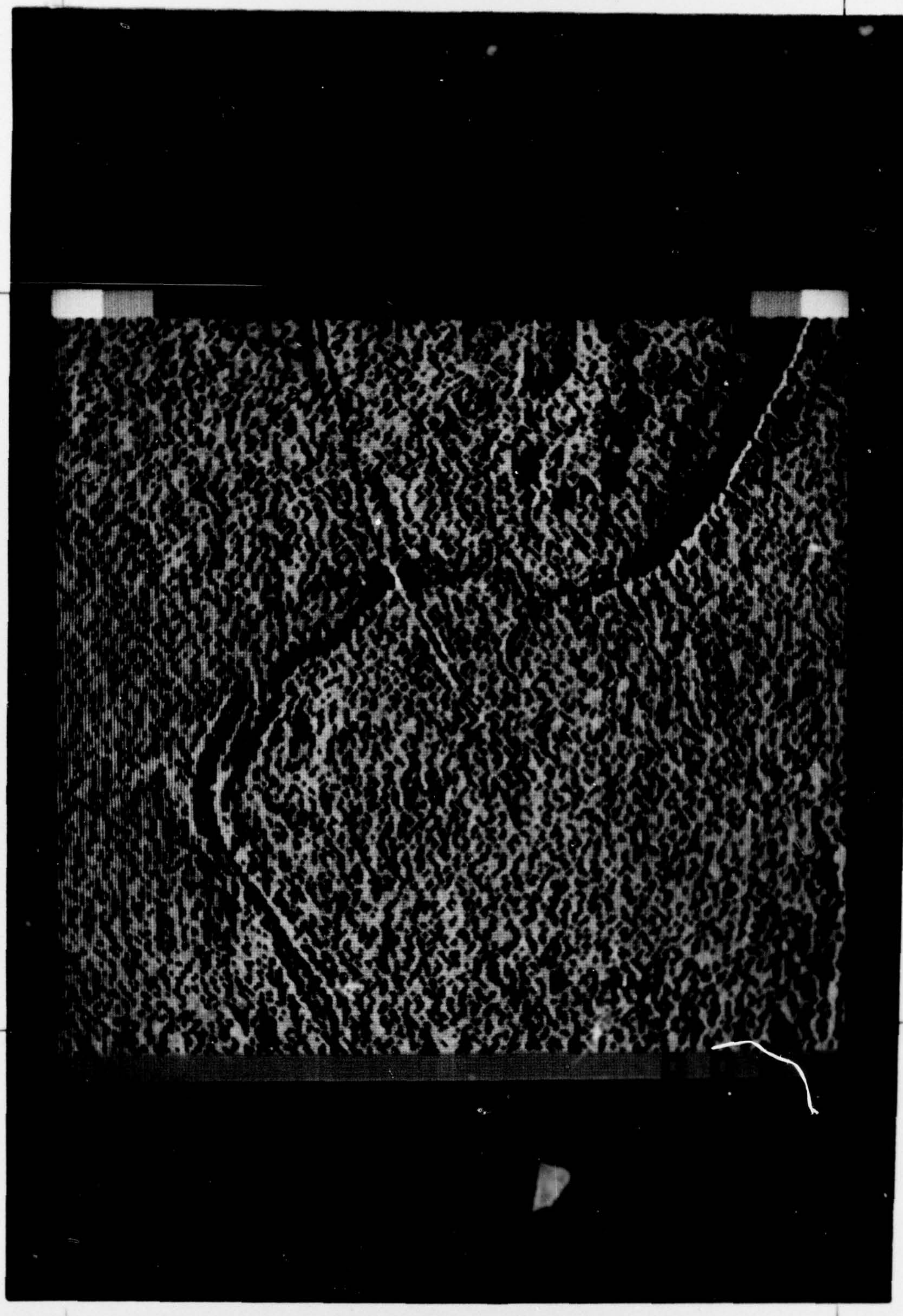
Figure 26. Generic multilobe filter characteristic.

36



F. 27d

to 79%



F. 27b

to 79%

1 2

Original Enhanced

3 4

Original Enhanced

Figure 27. One-dimensional Fourier multilobe filter enhancement.

degradation of the images compared to the originals, this filtering technique does not warrant further investigation.

One-Dimensional Fourier Multinotch Filter

The disappointing results obtained with the multilobe filter indicated that a different filter characteristic representing a different filtering philosophy was in order. The concept of "tuning" the filter to a specific target size or spatial extent still was thought to be desirable. The multinotch filter incorporates this property, as did the multilobe filter. Given the target spatial frequency amplitude spectrum

$$\frac{\sin \frac{\pi f n}{N}}{\frac{\pi f n}{N}},$$

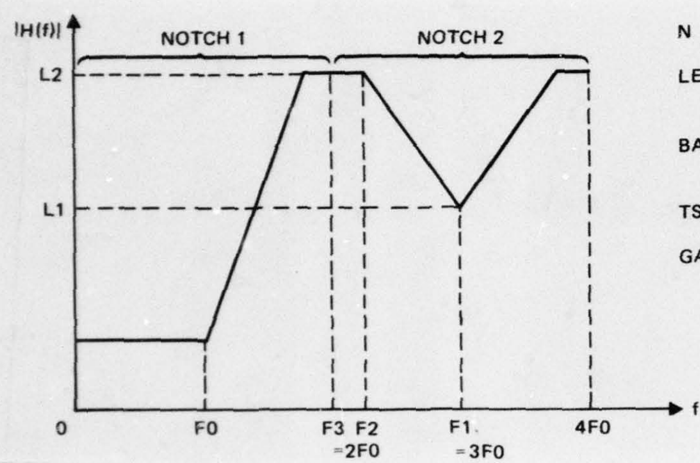
an inverse filter with frequency characteristic

$$\frac{\frac{\pi f n}{N}}{\sin \frac{\pi f n}{N}}$$

would have the property of emphasizing high spatial frequency components of a target signature of a certain spatial extent. The difficulty with this inverse filter is that its gain tends to infinity for spatial frequencies near the values of $f = kN/n$, $k = 1, 2, 3, \ldots\ldots$, because the $\sin \pi fn/N$ term in the denominator goes to zero at these points. Thus an approximation to the inverse filter must be implemented.

The approximations employed were to limit the filter gain to not more than a specified value over the applicable range of spatial frequencies and to use a piecewise linear function as the filter characteristic. Figure 28 shows a generic multinotch filter characteristic. As with the multilobe filter, the number of notches N is half of TSIZE, which is the spatial extent in pixels of the target of interest. Two notches are shown in Figure 28; hence, the corresponding target size is 4 pixels. The target extent TSIZE determines the spatial frequency location at which the filter gain first starts to decrease. For TSIZE = 4, FO = 32 cycles per image dimension. F. 28

The parameter BAND specifies the width of the notches at their top as a fraction of 2FO, which is the maximum possible value; BAND takes on



TA/3783

N = NUMBER OF LOBES

LEVEL = $\frac{L2}{L1}$, 1 < LEVEL

BAND = $\frac{F1 - F2}{F1 - F3} = \frac{3F0 - F2}{F0}$, 0 < BAND ≤ 1

TSIZE = $\frac{128}{F0}$ = 2N

GAIN = AMPLITUDE SCALE FACTOR

Figure 28. Generic multinotch filter characteristic.

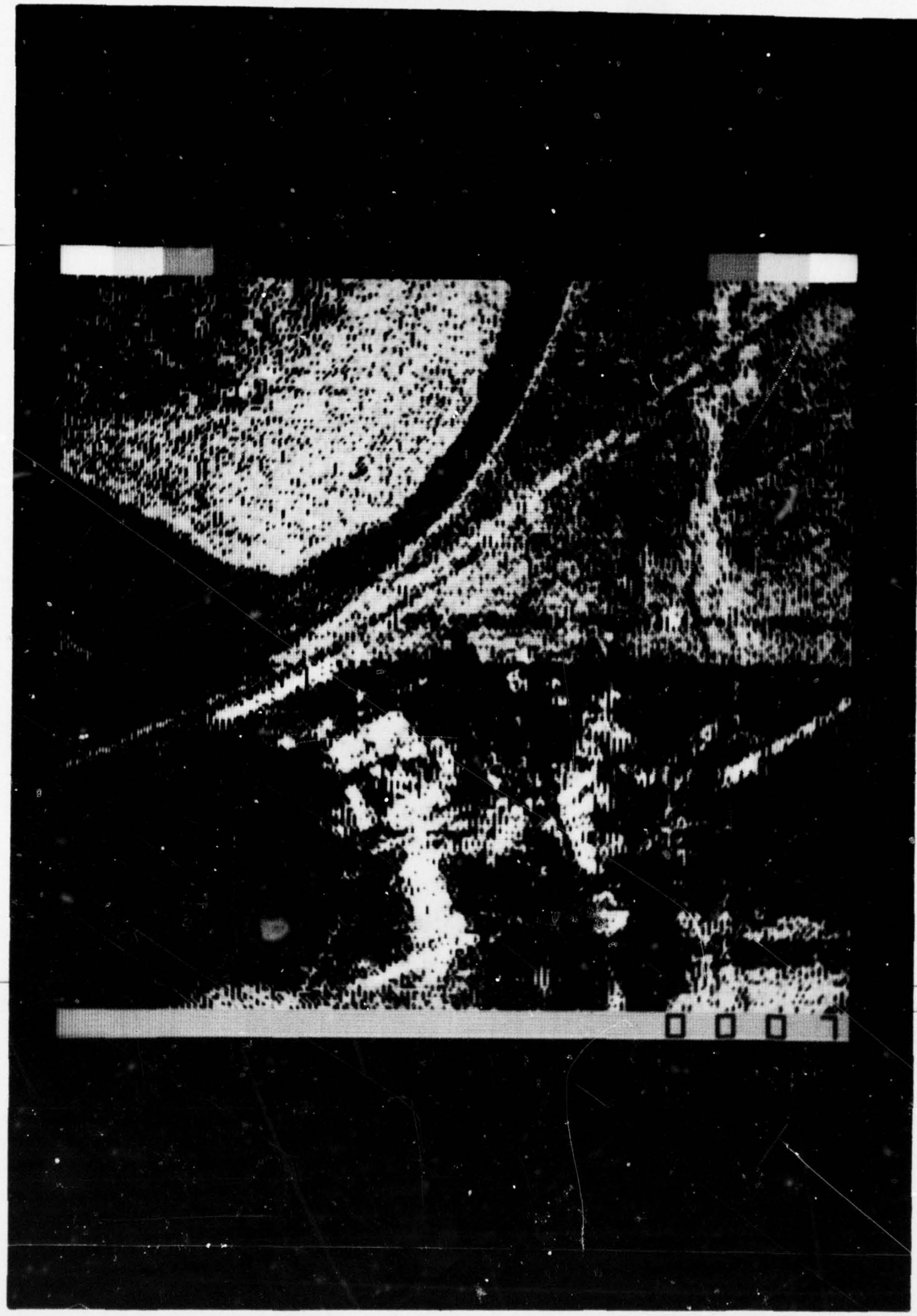
39

values between 0 and 1. The parameter LEVEL is the ratio of the maximum allowable relative filter gain to the local minimum of the relative filter gain in the last notch. The parameter GAIN is a multiplicative amplitude scale factor on the entire filter characteristic. The parameters TSIZE, GAIN, BAND, and LEVEL were independently variable.

The optimization of the filter performance with respect to the parameters GAIN, BAND, and LEVEL produced similar results to those obtained with the multilobe filter. A value of BAND = 1 gave the best performance, and a value of LEVEL = 2 resulted in a sufficiently high value of maximum relative filter gain. Values of GAIN on the order of 10^{-3} gave the best results. The sequence of operations used to apply the multinotch filter was the same as that for the multilobe filter previously described. Only the filter characteristic was changed. The filtering direction was horizontal. Figure 29 shows multinotch filtered images obtained with parameter values GAIN = 0.001, TSIZE = 2, BAND = 1, and LEVEL = 2. Fig 29

Since the inverse filter tends to make the amplitude spectrum of targets of spatial extent TSIZE = 2 approach that of an impulse function, the effect on the filtered image is that the target-sized objects on a line in the filtering direction appear bright and everything else on that line appears dark. This is particularly evident in the bridge scene by the dark horizontal line which crosses the bright pixels corresponding to the bridge over the river and in the train scene where the high intensity returns from the train and power plant contain two-pixel-sized scintillations which cause a large filter output at those points. Because of the logarithmic truncation scheme, intensity distributions which are bimodal with peaks in the bright and in the dark intensities will result in an image which appears to be binary. Due to the "tuning" of the multinotch filter to a specific spatial frequency and the fact that it shapes the amplitude spectrum of its response to a target toward that of an impulse function, the multinotch filter tends to produce an output intensity distribution which is bimodal. The distribution peak in the bright intensities reflects the filter's large response to target-sized objects; the distribution peak in the dark intensities reflects its small response to non-target-sized objects. This bimodality of filter response, coupled with the fact that a radar image of a relatively homogeneous ground area of large spatial extent contains many return intensity variations of small spatial extent within it, results in many areas in the image which cause a large filter

40



F. 29d

407920

F. 29b

to 79%

a 1

b 2

Original

Enhanced

c 3

d 4

Original

Enhanced

Figure 29. One-dimensional Fourier multinotch filter enhancement.

41

response. This in turn reduces the target-to-background contrast and generally makes *tactical-sized targets more difficult to detect.*

Two-Dimensional Fourier Modified Crispening Filter

For reasons of computational simplification and lower processing costs, our initial investigations of transform domain filtering techniques were confined to one-dimensional processing. To answer the question of whether filtering in both image spatial dimensions would improve the results with the transform domain techniques, the one-dimensional technique which is considered to be the most promising, namely the Fourier modified crispening filter, was extended to a two-dimensional process.

The input SAR images had 256 resolution elements in each spatial dimension. The 65,536-word array which is required to store the entire image in floating point format would not fit in the available amount of Sigma 5 computer memory. Thus, to perform two-dimensional filtering on this array using a 256-point Fourier transform required a large number of disk input-output operations to achieve the necessary image transposition. Image transposition in this fashion is often referred to as "corner-turning". However one quadrant of an image, a 128-by 128-pixel sub-image, would fit in the available amount of computer memory. This leads to processing each 256 by 256 image by using a 128-point Fourier transform on each of the four image quadrants of 128 by 128 pixels. Each quadrant could be processed separately and later reassembled into a single image. The advantage of using a 128-point transform over the 256-point transform was that fewer disk input-output operations would be required, and therefore processing time and expense would be reduced. Such a process was implemented.

The program read the original floating point data from the disk, one row of 128 pixels in the first quadrant at a time. The 128-element row of pixels was then Fourier transformed, multiplied by a set of 128 filter coefficients, and inverse Fourier transformed. The filtered data were then multiplied by the weighting factor WFACT, subtracted from the original data, and the result was stored in a row of a 128 by 128 buffer array. Each row of the quadrant was processed in this manner until one complete quadrant had been read in, processed in one dimension (rows), and had filled the buffer array.

Each of the 128-element columns of the buffer array were then processed as follows. The column was transferred to a temporary storage array which was Fourier transformed, multiplied by the same set of 128 filter coefficients as applied to the transformed rows, and inverse Fourier transformed. The filtered column data were then multiplied by the weighting factor WFACT, subtracted from the corresponding column of row-processed data in the buffer array, and the difference was stored in the correct column in the buffer array. When all 128 columns in the buffer had been processed, the data in the buffer array were logarithmically converted from floating point to fixed point mode and output to a temporary storage file on the disk. The entire process was then repeated on the second, third, and fourth quadrants of the original image, and each quadrant was output to a unique temporary storage file on the disk.

When all four quadrants had been processed, the temporary files of processed data were read back into the computer memory and reassembled into a single image. The final processed image of 256 by 256 pixels was assembled one row at a time in memory and output one row at a time to a permanent storage file on the disk, from which it later was displayed.

It was anticipated that the form of two-dimensional filtering described above would produce some boundary effects at the points where different quadrants butted against each other. Although such effects are not desirable, the advantage in input-output efficiency was deemed to outweigh any disadvantage of these boundary effects, particularly for an initial study. To make it possible to isolate unwanted boundary effects due to the contrast enhancement properties of the filter, the floating point original data on the disk were processed as described above with a filter weighting factor of zero. This produced new originals which show quadrant boundary effects.

The generic filter characteristic used in two-dimensional Fourier modified crispening is the same as that for the one-dimensional modified crispening filter shown in Figure 24, and the remarks concerning optimization of the filter performance with respect to the filter parameter values apply here as well. Figure 30 shows the two images processed by two-dimensional modified crispening. The filter parameters, scaled to an image with 128 pixels per dimension, were GAIN = 0.002, FREQ = 40, BAND = 24, LEVEL = 0, and WFACT = 0.25.

Fig 31

A

B

Original

Enhanced

C

D

Original

Enhanced

Figure 30. Two-dimensional Fourier modified crispening enhancement.

It is apparent that two-dimensional modified crispening improves target-to-background contrast, particularly for isolated targets such as the convoy at the lower right of the bridge scene or the pair of objects in the upper left corner of the train scene. The strong returns from the power plant complex produce some intensity smearing. This effect can be minimized by thresholding prior to filtering. Disregarding the effects of quadrant processing in Figure 30 and comparing these results to the one-dimensional modified crispening filter (Figure 25), the two-dimensional version clearly has a greater effect on image contrast. However, the higher contrast observed in Figure 30 is not judged to provide a sufficient operational performance increment to be cost-effective. A hardware mechanization would have to implement the functional equivalent of the process and would neither be simple nor inexpensive. In view of these considerations, further efforts in this type of two-dimensional filtering, particularly for real-time applications, should perhaps be directed toward radar or other sensors with higher resolution than the 20-foot resolution SAR data used in this program.

Two-Dimensional Fourier Matched Filtering

Having implemented a form of two-dimensional filtering, it was decided to briefly investigate matched filtering as a technique for better utilization of dynamic range. Matched filtering is a technique which comes from signal detection theory (Van Trees, 1968; North, 1963; Nahi, 1969) in which the optimum linear filter for detecting a known signal in the presence of noise is the filter whose transfer function in the Fourier domain is the same as that of the known signal which is to be detected. This filter is optimum with respect to the criterion of maximizing the detected signal to filtered noise ratio at the output of the filter.

Although the target signatures of interest are two-dimensional and therefore have two-dimensional Fourier transforms, these two-dimensional transforms are separable into a product of two one-dimensional transforms, each corresponding to one of the orthogonal SAR sampling directions. Thus, the two-dimensional Fourier transform of the rectangular target can be written as

$$\frac{\sin \frac{\pi f_x n}{N}}{\frac{\pi f_x n}{N}} \; \frac{\sin \frac{\pi f_y m}{N}}{\frac{\pi f_y m}{N}} ,$$

46

where f_x and f_y represent spatial frequency and n and m represent target size in the x and y directions, respectively. This is the transfer function of the matched filter for detecting a target in noise. Under this filtering philosophy, noise is not only radar scintillation, spurious signals in electronic components induced by thermal effects, and processor roundoff and truncation errors, all of which are commonly considered sources of noise, but also any image information produced by objects in the scene which are not of the size we seek to detect. However, image information other than target signature is important in SAR target detection for orientation and context cues. Thus, a useful filter must not eliminate too much non-target information.

The matched filtering evaluation was restricted to targets which were square. This allowed some simplification in the processing program because the same filter characteristic could then be employed in both directions. The generic filter characteristic for one filtering direction is shown in Figure 31. Since the target was assumed to be square, the filter characteristic for the other direction was identical.

The overall two-dimensional matched filter has a characteristic which is a surface in three dimensions, a cross-section of which, taken at a constant spatial frequency in one dimension, has the shape shown in Figure 31. The number of lobes in one dimension is half the target spatial extent TSIZE, in pixels, in that dimension. The parameter GAIN is an overall amplitude scale factor and from the standpoint of matched filter theory has no effect on the output signal-to-noise ratio, so it can be set to 1.

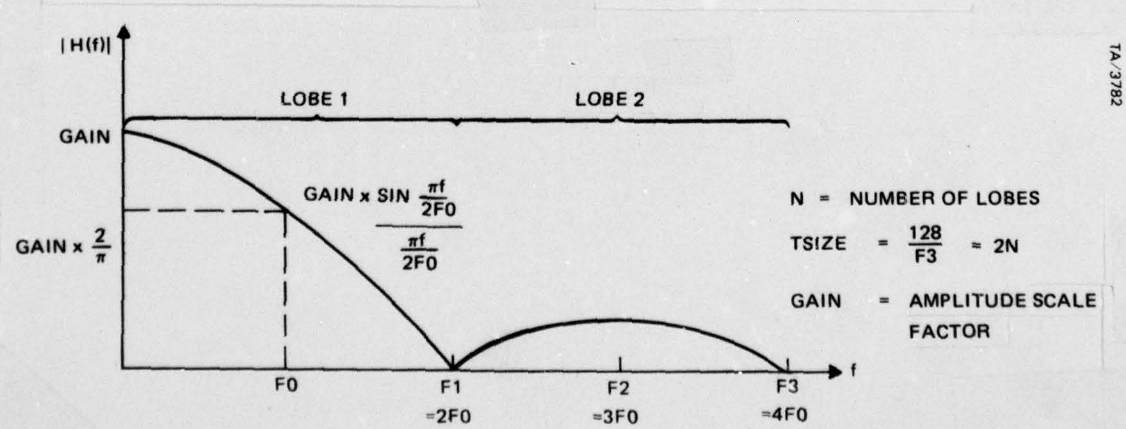


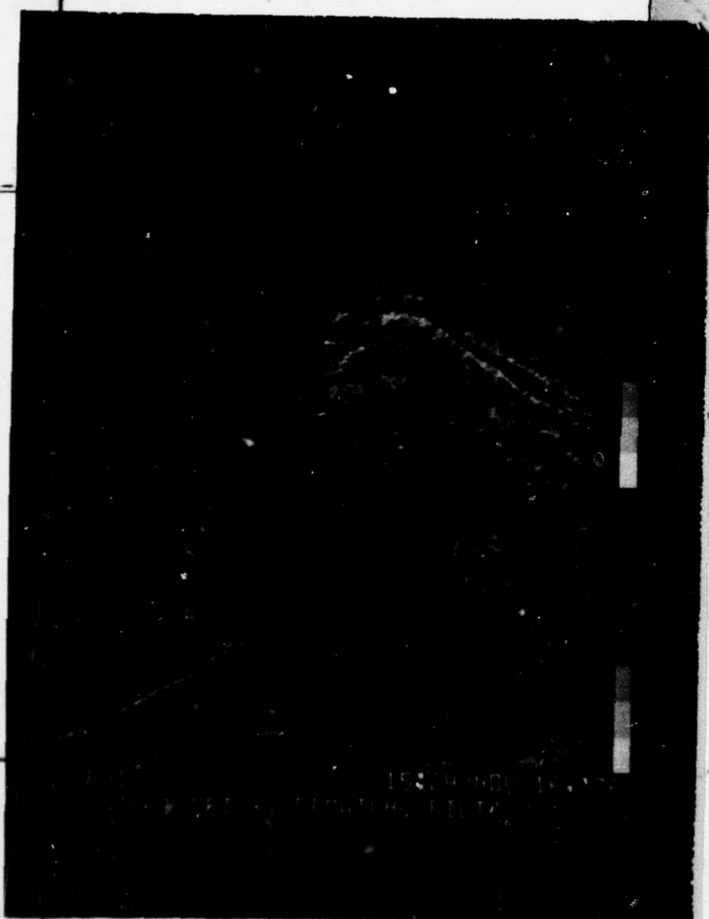
Figure 31. Generic matched filter characteristic.

The filtering procedure was as follows. The first quadrant of floating point data was read row by row from the disk and each row was filtered in one dimension. The one-dimensionally filtered data were then multiplied by a weighting factor WFACT, added to the original data, and the sum was stored in a buffer array. The weighting factor WFACT was used to control the relative amounts of original and filtered data comprising the final processed image to prevent the elimination of too much non-target information. Having filtered each row of the first quadrant, each column was filtered with the same transfer function as applied to the rows. The filtered columns were multiplied by WFACT, added to the corresponding column of row-processed data in the buffer array, and the sum was stored in the appropriate column of the buffer array. The ratio of original data to two-dimensionally filtered data in the final image was therefore 1 to $(WFACT)^2$. After all the columns in the first quadrant were processed, the data in the buffer array were logarithmically converted to fixed point mode and output to a temporary storage file on the disk. The process was then repeated on the second, third, and fourth quadrants, and the final image was reassembled from the processed quadrants as described previously.

Figure 32 shows the results of two-dimensional matched filtering with TSIZE = 2 pixels, i.e., objects which were 2 x 2 pixels (40 x 40 feet) in size. So as not to lose context information, a conservative weighting factor of WFACT = 0.25 was used, which means that the filtered scene data had a relative weight of 1/16 that of the original scene data. At this low relative weighting, an improvement in small target contrast can be discerned, as exemplified by the convoy at the lower right of the bridge scene. However, intensity smearing from the strong return areas occurs and is present in the train scene. This can be largely eliminated by thresholding prior to filtering. The filtering program was modified to perform the threshold operation.

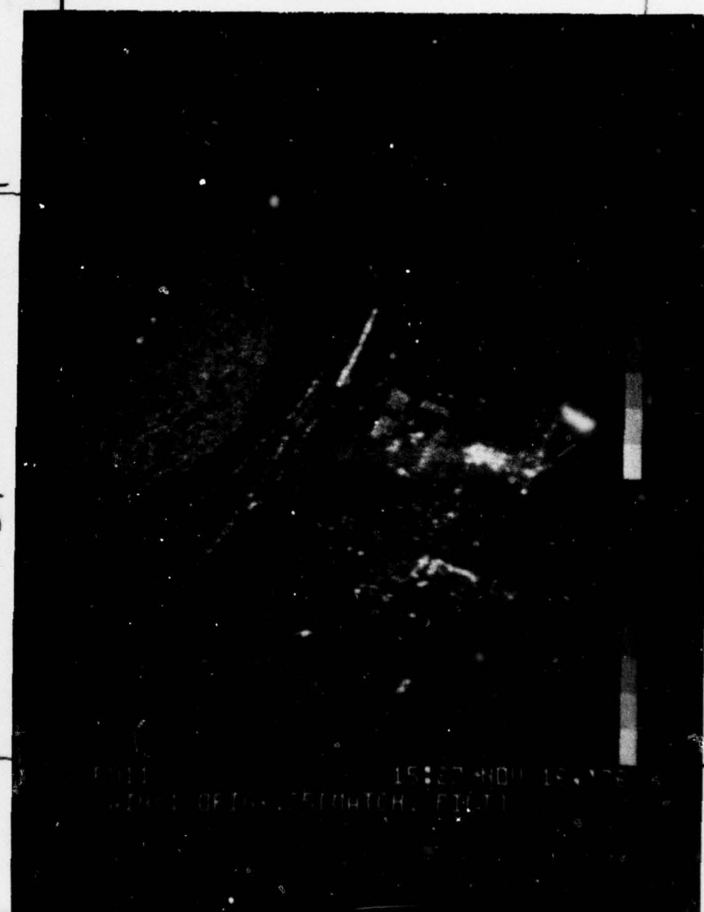
As each row of floating point data was read from the disk, the pixel intensity values were compared to a threshold. Those pixels whose intensity value exceeded the threshold had their intensity value replaced by the threshold value. After this operation, the filtering process went through as previously described. The threshold value was a number which represented a given percentile point on the image intensity distribution, as determined from SAR image histograms and cumulative distributions.

TOP



F. 32B

TOP



To 1270w

F. 32D

1

2

Original

Enhanced

3

4

Original

Enhanced

Figure 32. Two-dimensional Fourier matched filtering enhancement.

48

The 95th percentile of the image intensity distribution was found to produce good suppression of intensity smearing, as shown in Figure 33. The weighting factor used in Figure 33 was WFACT = 2.5, so the filtered data had a relative weight of 6.25 times the original data. The filter characteristic was the same as that used for the non-threshold images in Figure 32. The effects on small target contrast with thresholding are about the same as without thresholding. The thresholded images in Figure 33 are better due to the lack of smearing. Some loss of context information seems to have occurred in the thresholded images. This was due to the higher relative weighting of the filtered as opposed to the original data. The effect is as expected and indicates that relative weighting was too high in favor of the filtered data.

As was the case with the two-dimensional Fourier modified crispening filter, the two-dimensional Fourier matched filter enhanced small target contrast. The overall improvement of SAR image quality for operator performance improvement, however, was judged to be insufficient to justify further examination.

CONCLUSIONS AND RECOMMENDATIONS

Eleven monochromatic image enhancement techniques were investigated for their potential to provide improved SAR operator tactical target detection performance. Existing and new filter characteristics were proposed and implemented. Evaluation of the resulting processed images confirmed that the techniques had the intended effect on the image data. In most cases, the observed effect was judged desirable, but not of sufficient magnitude to yield any significant increase in operator target detection performance. The two exceptions to the preceding statement are multilobe and multinotch filtering. In these instances, the observed effect was so large as to be undesirable. So much contextual information had been filtered out that the processed image was less useful for target detection than the original image. The problems with multilobe and multinotch filtering could be alleviated by using a weighted sum of the original data and the filtered data, as was done with the two-dimensional matched filter.

In view of the results obtained with the one-dimensional filters and the two-dimensional matched filters, we do not expect the monochromatic enhancement techniques such as investigated here to be significantly better

1

2

Original

Enhanced

3

4

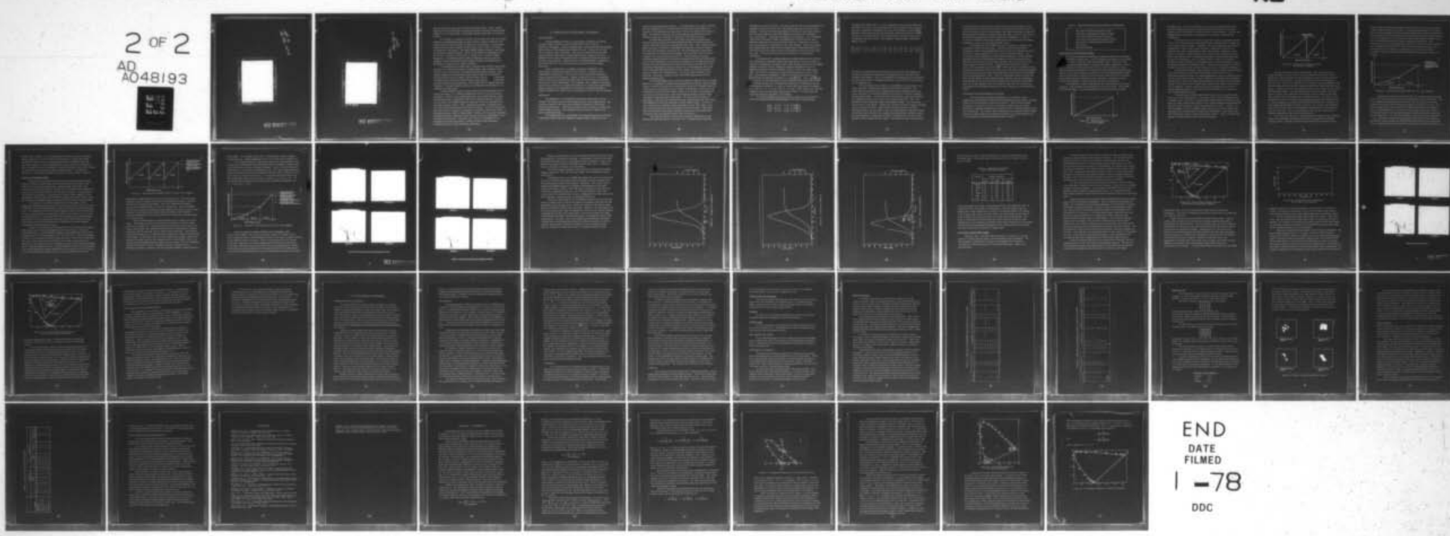
Original

Enhanced

Figure 33. Two-dimensional Fourier matched filtering with thresholding.

AD-A048 193 HUGHES AIRCRAFT CO CULVER CITY CALIF DISPLAY SYSTEMS--ETC F/G 17/9
SYNTHETIC APERTURE RADAR IMAGE PROCESSING TECHNIQUES DEVELOPMEN--ETC(U)
OCT 77 E J DRAGAVON, M L HERSHBERGER F33615-76-C-1115
UNCLASSIFIED HAC-P77-189 AFAL-TR-77-113 NL

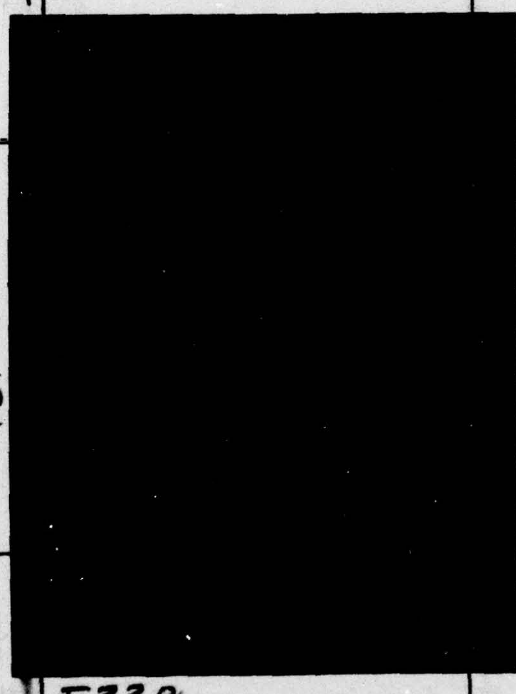
2 OF 2
AD
A048193



END
DATE
FILMED
1 -78
DDC

Printer: Top left to right

TOP

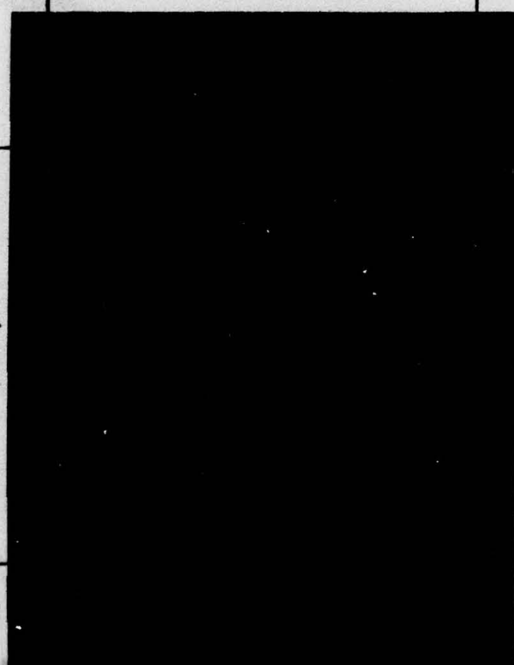


+0127w

F.33B

Printer: Flop left to right

TOP



TO 127%

F. 33D

than the original images with regard to operational utility. Thus, further exploration of monochromatic processing techniques with high resolution SAR imagery is not recommended. Ultra-high resolution SAR may be a different matter.

In retrospect, a problem common to all the black and white processing techniques investigated was that the small tactical targets of interest in the image occupied a limited number of pixels in each dimension. For example, with the 20-foot radar resolution of the images used, a tank or truck only occupied one or two pixels. This small spatial extent resulted in wide extent in the spatial frequency domain. Hence, it was difficult to distinguish targets from non-targets by attempting to separate their respective spatial frequencies. The difficulty was not one of making the filters perform as designed, but was associated with the spatial frequency characteristics of the targets of interest in the images provided.

A worthwhile effort might be to re-evaluate some of the 11 processing techniques, or improved versions, on SAR imagery with resolution on the order of one or two feet and correspondingly less ground coverage. Small tactical targets such as tanks and trucks would occupy 10 or 20 pixels in each dimension in such an image, and the smaller coverage would mean that 10 pixels would represent a larger fraction of the total number of resolution elements along an image dimension. This implies that the spatial frequency "signature" of tactical targets would be more distinctive and perhaps more amenable to enhancement.

Another possible avenue for future efforts is the investigation of cascaded processing, i.e., sequential application of two or more processing techniques. If each of two cascaded processes can be described by a linear filtering operation, a single linear filter whose spatial frequency response is the product of the spatial frequency responses of the two cascaded filters will yield identical results. Thus, cascaded processing is sensible only if one of the processing techniques involved is nonlinear. Several of the techniques discussed are nonlinear, and nonlinear versions of some of the linear filtering processes could be implemented. However, the comments regarding the spatial frequency signatures of the targets of interest would apply to cascaded processes as well. Therefore, the overall recommendation is to investigate cascaded processing or refined versions of the techniques evaluated here only on SAR with higher resolution.

3.0 PSEUDOCOLOR ENHANCEMENT TECHNIQUES

BACKGROUND

Another potential means of obtaining better utilization of the wide dynamic range of synthetic aperture radar (SAR) video is the use pseudocolor encoding. The term "pseudocolor" refers to the fact that original SAR images are two-dimensional intensity records. In other words, they have no intrinsic color properties. One class of methods which map SAR sensor dynamic range onto a limited display dynamic range is the use of additional information encoding dimensions to expand the effective amount of dynamic range available. Because of the color discrimination capability of the human observer, color was selected as a potentially useful display dimension.

Color comprises three variables: luminance (brightness), hue (wavelength), and saturation (purity); each of which, in theory, can be independently manipulated. Thus color provides a great deal of potential image transformation flexibility. Another reason for the use of pseudocolor SAR encoding is the success with which color has been used as an encoding dimension in similar applications. For example, color coding has long been used to depict political or topographic boundaries in cartographic applications. Another example is the application of pseudocolor encoding techniques to imagery obtained from earth resources satellites as an aid to land use analysis.

APPROACH

The equipment used to generate pseudocolor SAR images was the Aerojet color display and associated image memory, color map memory, and computer software at the University of Southern California (USC) Image Processing Laboratory. The Aerojet system is the forerunner of present day Comtal display systems. This equipment and the manner in which it was used are described below.

The original SAR gray shade image to be operated upon was read from a standard 9-track computer compatible magnetic tape in digital form and stored in the display image memory. The image memory was used to refresh the color display CRT at a 30-Hz frame rate and could store images up to 512 lines of 512 pixels per line with up to 6 bits of gray shade information per pixel. To maintain continuity with the monochrome processing work, the original images were restricted to 256 lines of 256 pixels per line.

As described in Section 2, all 21 files of log filter magnitude data underwent preprocessing which exponentiated the data and performed the 4:1 overlay required to obtain an actual ground map image. The resulting FLAMR SAR image files of 348 lines of 384 pixels per line were stored on the disk with each pixel intensity represented as a 32-bit floating point real number. In preparation for pseudocolor processing, the image files on the disk were cropped to 256 lines of 256 pixels per line. These cropped sub-images underwent postprocessing to convert the representation of the pixel intensities from 32-bit floating point real numbers to integers in 8-bit fixed point format. Finally, since the image memory of the Aerojet system provided only 6 bits of intensity information storage per pixel, the 6 most significant bits of the 8 bits remaining after postprocessing were written on a 9-track magnetic tape to be input to the Aerojet system. Each gray shade image file on the final tape, therefore, consisted of 256 lines of 256 pixels per line with 6 intensity bits per pixel.

The Aerojet system has a color map memory, separate from the image memory, in which up to 64 different color maps can be stored. The particular map to be used at any given time is selected by means of a trackball controller which moves a map selection cursor symbol on the display. Each color map is specified by 64 triples of numbers. Each of the 64 triples corresponds to one of the 64 possible values of the 6-bit fixed point binary numbers which represent image intensities at each pixel. The values of the numbers in each triple, referred to as tristimulus values, determine the hue, saturation, and brightness of the color assigned to a given shade of gray. The Aerojet color map memory provided 12 bits of storage for each color triple, which was 4 bits for each of the three numerical values in the triple. This meant that the numerical values used to specify the amounts of the various color primaries had to be integers between 0 and 15 inclusive.

The numerical values of each color triple were expressed in octal format and were punched on a computer card. A complete color map was specified by a deck of 64 cards. The deck could be read into the computer and *assigned to any one of the 64 color map storage locations in the color map* memory. As the gray shade image data were read out of the image memory, the 6-bit pixel value was used as an address to a location in the selected color map memory. This location contained a 12-bit word. The 4-bit digital values found in the left, middle, and right thirds of this word were each digital-to-analog (D/A) converted and the resulting voltages drove the blue, red, and green guns, respectively, in the color display. Hence, the original gray shade image was seen with each distinct shade of gray portrayed by the color specified by the values on the corresponding card in the selected color map deck. A discussion of the theory of colorimetry is presented in Appendix A.

The pseudocolor encoding maps were initially defined in the 1960 CIE UCS chromaticity coordinate system (See Appendix A), because the general objective of pseudocolor encoding was to obtain equally perceptable differences between successive colors, analogous to equally perceptable luminance differences between successive shades of gray in a black and white image. Thus, each color in a pseudocolor map was first specified in terms of its (u, v) chromaticity coordinates. These were then transformed to (x, y) chromaticity coordinates according to $x = \frac{3u}{2u - 8v + 4}$ and $y = \frac{2v}{2u - 8v + 4}$. This was done to take advantage of the brightness matching properties of the XYZ primary system and to simplify the computation of tristimulus values. The tristimulus values X, Y, and Z of a color with chromaticity coordinates x, y, and $z = 1 - x - y$ are $X = \frac{x}{y} B$, $Y = B$, and $Z = \frac{z}{y} B$, where B is a desired luminance value for that color, normalized so that $B = 1.0$ represents the maximum luminance capability of the display.

The XYZ tristimulus values were then transformed to the RGB tristimulus coordinate system of the display by the linear transformation

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 2.99 & -1.34 & -0.490 \\ -1.06 & 1.97 & 0.056 \\ 0.111 & -0.260 & 1.03 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

325

as specified in Wallis (1975). To precompensate for film nonlinearities and chromatic distortion the RGB tristimulus values obtained in the preceding step were nonlinearly transformed to a new set of tristimulus values $\overline{R}\overline{G}\overline{B}$. Exposure of the color film with a color whose tristimulus values were $\overline{R}\overline{G}\overline{B}$ resulted in the desired color on the processed film, namely a color whose tristimulus values were RGB. The nonlinear transformation, obtained by Wallis (1975) is

$$\begin{bmatrix} \overline{R} \\ \overline{G} \\ \overline{B} \end{bmatrix} = \begin{bmatrix} 2.982 & -1.369 & 0.113 & -1.654 & 0.771 & -0.580 & 0.211 & 0.082 & -0.514 & 0.897 & -0.291 & 0.171 \\ -0.136 & 1.744 & 0.178 & -0.302 & -1.276 & -0.261 & -0.004 & -0.038 & 0.040 & -0.122 & 0.616 & 0.140 \\ 0.886 & -1.480 & 2.453 & -1.259 & 0.710 & -1.903 & -0.008 & 0.015 & 0.453 & 0.454 & -0.406 & 0.894 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \\ R^2 \\ G^2 \\ B^2 \\ GB \\ RB \\ RG \\ R^3 \\ G^3 \\ B^3 \end{bmatrix}.$$

The step of compensating the RGB tristimulus values for film effects was omitted in preparing the pseudocolor mapping decks to be used for direct viewing on the display. Thus, two pseudocolor mapping decks were prepared for each version of each technique evaluated — one for direct viewing and one for recording on film.

The precompensated tristimulus values $\overline{R}\overline{G}\overline{B}$ were converted to 8-bit binary numbers, representing integer values between 0 and 255. The conversion was performed by multiplying each precompensated tristimulus value, nominally a real number between zero and one, by 255 and rounding to the nearest integer. A check was made at this point to determine whether the resulting number was less than zero or greater than 255. If so, this indicated a color outside the range of the display system. The reason why this can occur is as follows. Assume the intent is to produce a highly saturated red on the processed film. Since the film recording process results in saturation losses, due primarily to the overlapping spectral absorption bands of the dyes used in the film, the nonlinear precompensation transformation of the RGB tristimulus values has the effect of increasing the saturation of the displayed color to offset the film losses. The display system has limitations on the saturation it can achieve for any given hue, and this implies an even lower limit on the saturation which can be realized on the processed film. If any of the desired colors in a pseudocolor map under consideration required precompensated tristimulus values outside the range of 0 to 255, all colors in the map were appropriately scaled in saturation, brightness, or both, to maintain the desired chromaticity relationships until all the precompensated tristimulus values were within the allowable range.

The tristimulus values $\bar{R}\bar{G}\bar{B}$ were then corrected for nonlinear CRT brightness transfer function effects. This was necessary because the tristimulus values of the displayed color on the CRT were nonlinear functions of the drive signals applied to the red, green, and blue phosphors. The correction was accomplished by a table look-up technique. Each of the 256 possible tristimulus values defined an address in the look-up table and at that address was found the corrected tristimulus value, which was an 8-bit binary integer between 0 and 255. The same table values were used to correct all three tristimulus values $\bar{R}\bar{G}\bar{B}$. While this was not entirely correct, because the brightness transfer functions of the red, green, and blue phosphors were not exactly the same, they were sufficiently similar that the associated error was small (Wallis, 1975). The 8-bit tristimulus values read from the look-up table had the 4 least significant bits removed, and the remaining 4 most significant bits were used as the display input tristimulus values $\tilde{R}\tilde{G}\tilde{B}$. Only the 4 most significant bits were used, because the color map memory associated with the display allowed only 4 bits of storage for each tristimulus value.

PSEUDOCOLOR TECHNIQUES EVALUATED

Seven pseudocolor encoding techniques were evaluated, using independent and redundant information encoding dimensions to convey a greater portion of the SAR dynamic range to the operator. Some variation within techniques was achieved, however, the effect of dynamically varying the pseudocolor mapping parameters as functions of time or image statistics could not be assessed. The relative value of pseudocolor encoding for tactical target detection in SAR imagery was obtained from static scenes generated for the seven pseudocolor encoding techniques listed in Table 5.

Table 5. PSEUDOCOLOR IMAGE ENCODING TECHNIQUES

One-Color Continuous Intensity Coding
Two-Color Discontinuous Intensity Coding
Two-Color Continuous Intensity Coding
Three-Color Discontinuous Intensity Coding
Three-Color Continuous Intensity Coding
Hue Coding
Saturation Coding

One- and Two-Color Intensity Coding

The first and simplest pseudocolor mapping considered was intensity modulation within a single color. This technique was employed in the FLAMR system by virtue of the green P20 CRT phosphor used. Moreover, SAR image transparencies generated by Goodyear Aerospace Corporation, using a laser, appear to be better than their black and white equivalents. It was therefore decided to implement a similar one-color intensity-modulated mapping on a color CRT to determine whether such a less sophisticated mechanization would produce results as good as those of Goodyear and to investigate the possibility that the attribute of color contributes something to target detection performance, compared to an achromatic presentation.

The diagram in Figure 34 illustrates the characteristics of one-color mapping. This mapping was implemented in red and in green. The results were disappointing. The green mapping was not as good as the FLAMR

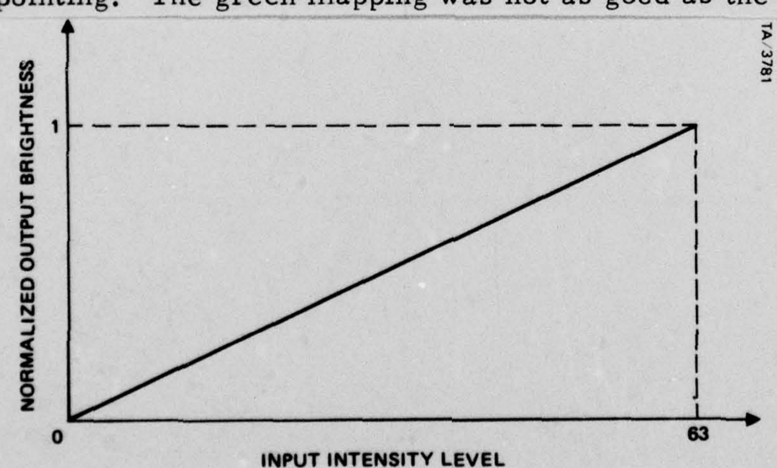


Figure 34. Intensity profile for one-color mapping.

transparencies. This is probably because the green FLAMR phosphor has a narrower spectral transmission than the spectral distribution of the green phosphor of the USC display, resulting in lower color saturation. The red also disappointing, particularly in comparision to a red laser generated SAR transparency.

The superiority of the red laser generated imagery is probably because a laser has a much narrower spectral width than a phosphor and is capable of much greater dynamic range. In addition, the spatial positioning and intensity modulation of a laser beam is more precisely controllable than that of an electron beam, which contributes to better color quality. Moreover, color CRT's are capable of higher dynamic range in black and white shades of gray than in any spectral color. In shadow mask types of color monitors (the type in used in this program), this is because as beam current is increased to increase the color luminance, the attendant spot size increase begins to excite neighboring phosphors of different colors. This causes saturation to decrease, moving the perceived color toward the white point on a chromaticity diagram.

In penetration phosphor color monitors, beam divergence with increasing beam current does not excite the other phosphor layers, but the use of a penetration phosphor monitor for imaging applications requires rapid switching of large acceleration voltages. This is difficult to accomplish with reliability and precision. Moreover, penetration phosphors are limited to two primaries, usually red and green, which significantly restricts the number of colors that can be displayed. Intensity modulated one-color mapping did not result in greater displayed dynamic range and was not considered to be a useful color CRT implementation for SAR.

Using two colors with the idea that the color of an image point would serve as a "most significant bit" of intensity code was the second pseudocolor technique tried. The color of a point would indicate whether it belonged to the upper or lower portion of the dynamic range, and its intensity would specify its position within the dynamic range slice assigned to its color. In theory, if a CRT is capable of 15 db of dynamic range in each of two colors, then by this technique 30 db of SAR dynamic range could be displayed — 10 db better than the 20 db dynamic range of typical monochrome CRT's. Figure 35 depicts the generic characteristics of such a two-color mapping.

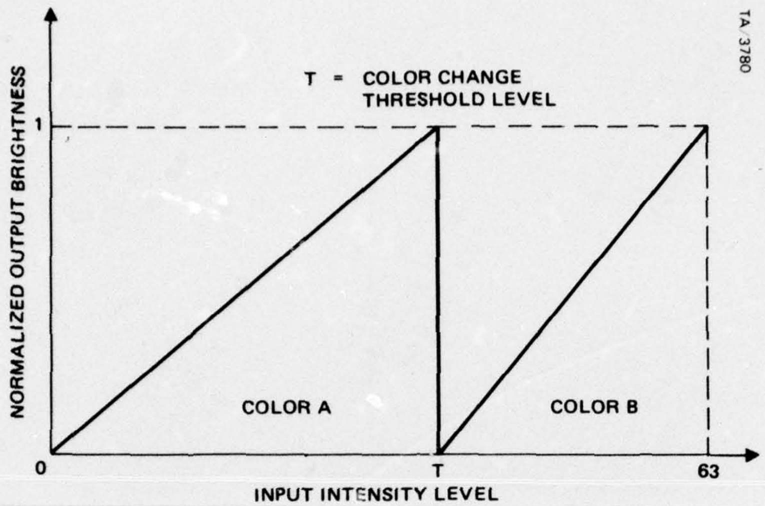


Figure 35. Discontinuous intensity profile for two-color mapping.

The two colors used were red and green. The value of the color change threshold T was 34, and the color assignments were red to color A and green to color B and the reverse assignment of green to color A and red to color B. The value of T = 34 out of a maximum of 63 is equivalent to a value of 34x4 = 136 out of a maximum of 255, and this value corresponds to the 97th percentile of the cumulative intensity distributions of the bridge and train scenes. Therefore only the very strongest returns were visible in color B with this threshold setting. Lowering the threshold to permit weaker returns to appear in color B was expected to allow more clutter to appear in the upper color band and was not thought to be a worthwhile pursuit. In addition it was noticed that because of the large intensity discontinuity at input intensity level T, returns appearing in color B at low brightness tended to be "swamped" by surrounding high brightness returns in color A. As a result, low brightness color B returns were not very conspicuous despite being of a different color than their surround. Due to these problems, the two-color discontinuous intensity mapping concept was not pursued further.

The use of two colors as a coding dimension redundant with brightness, rather than independent of brightness was the third pseudocolor mapping technique tried. Redundant coding has long been known to engineers as an efficient means of obtaining reliable (in the sense of error-free) transmission of information through a noisy channel. Human engineering data also support the concept of redundant information coding along multiple stimulus dimensions. The idea is that if the same information is coded along multiple stimulus dimensions, the probability of correctly perceiving the information is higher than if only one stimulus dimension is used for coding the information. The independence of color from brightness used in the previous two-color technique can be replaced by redundancy by requiring that the intensity profile of the redundant two-color mapping be continuous across the color transition threshold. Figure 36 shows the generic properties of this type of two-color mapping.

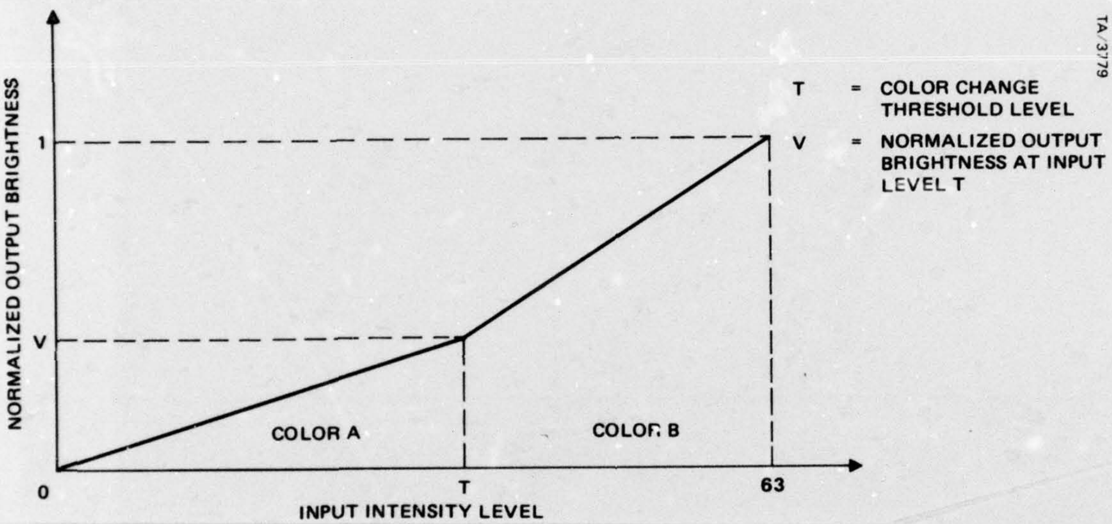


Figure 36. Continuous intensity profile for two-color mapping.

As with the previous two-color mapping, the colors used were red and green, with both of the two possible color assignments being investigated. The color transition threshold T was varied from 32 to 48, and the normalized intensity value at the color transition threshold varied from 1/3 to 1/2. The input intensity values used for the color transition threshold corresponded (after multiplication by four to account for the change in horizontal scale) to the 92nd percentile of the cumulative intensity distributions of the bridge and train scenes, for T = 32, and corresponded to nearly the 100th percentile for T = 48. As discussed under the discontinuous intensity two-color mapping, lowering the color transition threshold introduced clutter into the upper color band. A drop of only a few percentile points in the threshold was sufficient to produce such an effect. As with discontinuous intensity two-color encoding, target conspicuity was better with green assigned to the upper color band. Redundant coding, however, did not appear to have a beneficial effect with two-color mapping. It was reasoned that two colors may not be sufficient to add an adequate amount of redundancy or additional dynamic range to the coding process to produce significantly improved SAR images. It was decided, therefore, to investigate pseudocolor employing a larger number of colors.

Three-Color Intensity Coding

Based on our experience with the two-color intensity coded mappings, similar intensity coded mappings were developed with a third color, blue, added to the color stimulus dimension. Slight changes were made in the intensity profile characteristics to rectify some of the difficulties encountered at the color transition threshold in the two-color mappings. The addition of a third color was expected to 1) allow greater dynamic range capability in the independent coding technique and 2) allow an assignment of less dynamic range per color in the redundant coding technique to achieve better color saturation.

The changes in the intensity profiles were of two types. In the independent coding technique, the intensity did not drop all the way to zero after crossing a color transition threshold. Rather, the drop was to some residual level. The intention was that the residual brightness level would be better perceived in a high brightness background of another color than in a zero or very low brightness level. In the redundant coding technique, the intensity profile was allowed to have independently selectable slopes within each of the three color bands, subject to the constraints of continuity at the color transition thresholds and a finite dynamic range.

Figure 37 shows the generic characteristics of the three-color mapping with intensity and color as independent stimulus dimensions. Green was assigned to color C, representing the highest return intensities, because this assignment gave the best appearance in the two-color case. Blue was assigned to color A, representing the lowest return intensities, because it was expected that many of the lowest returns would come from water areas and such an assignment would exploit the "cool" perceptual property of blue (Kling and Riggs, 1971). Red was assigned to color B, representing the middle return intensities. The residual normalized output intensity level L was chosen to be 1/3, because this would allow the same total number of output intensity level and color

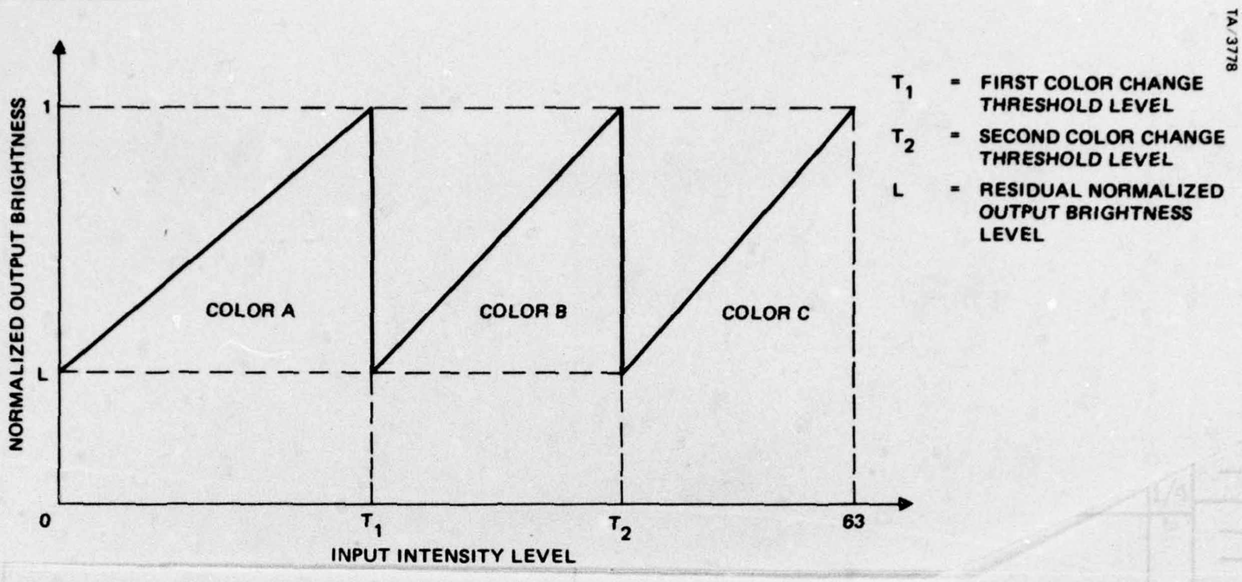


Figure 37. Discontinuous intensity profile for three-color mapping.

combinations (approximately 30) as in the two-color case. This also would hopefully reduce the effect of one color "overwhelming" another in intensity near the color transition thresholds. The color transition thresholds T_1 and T_2 were assigned values of 21 and 42, respectively. After multiplying by four to account for the difference in horizontal scale, these T_1 and T_2 were equivalent to the 24th and nearly the 100th percentiles of the cumulative intensity distribution for the bridge scene and to the 34th and 99th percentiles for the train scene.

Figure 38 illustrates the results obtained with the mapping described above. The bridge is easily seen in Figure 38, but the convoy at the lower right is not conspicuous. The interspersed red and blue in many areas of the image creates false color boundaries which tend to be interpreted as changes in terrain type and cause the image continuity to be disrupted. Lowering the upper color transition T_2 to make the convoy stand out in green would aggravate this problem. In the train scene, the train, power plant complex, and the two targets at the upper left are very conspicuous in the green band. Because of the presence of a significant water area in the train scene, the color boundaries between red and blue are more consistent with their interpretation as a land-water boundary. However, the false boundary problem is still present to some extent. The problem of one color "overwhelming" another of lower intensity has been largely eliminated by the use of a non-zero residual level.

The misleading nature of false image boundaries introduced by the use of color as an independent coding dimension lead to a redundant coding version of the three-color intensity modulated mapping. Figure 39 shows the generic characteristics of this type of pseudocolor map. The color assignments were the same as those used in the discontinuous intensity version — blue to color A, red to color B, and green to color C. The color transition thresholds were also maintained at their previous values, namely $T_1 = 21$ and $T_2 = 42$. The normalized output brightness levels at the color transition thresholds were $L_1 = 0.20$ and $L_2 = 0.50$.

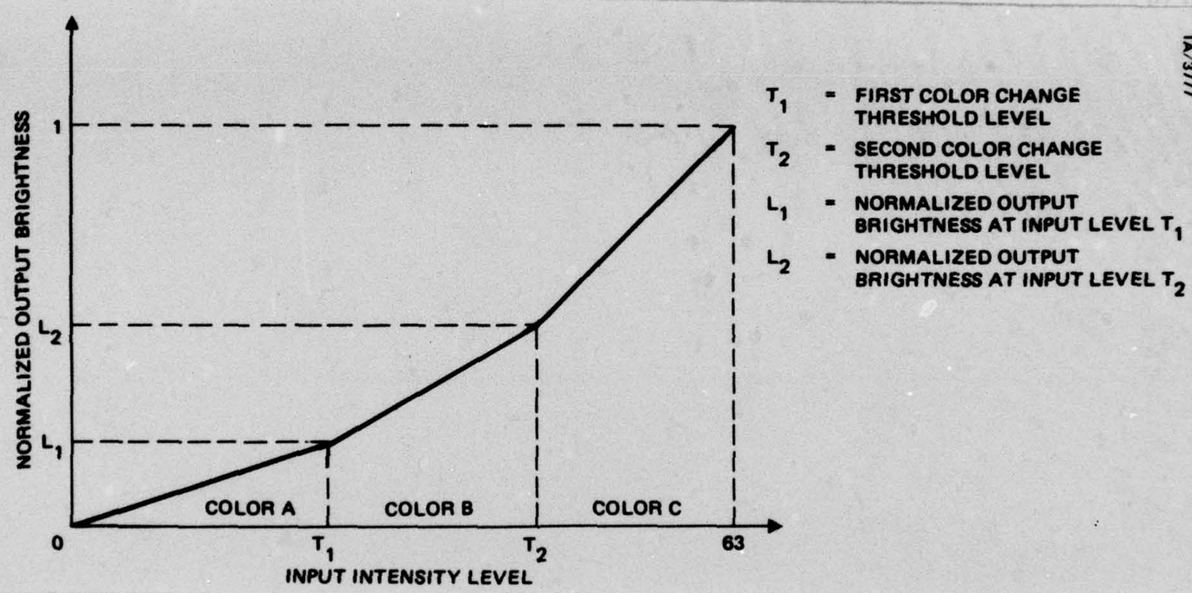


Figure 39. Continuous intensity profile for three-color mapping.

Figure 40 shows the images obtained with redundantly coded three-color mapping. The convoy at the lower right of the bridge scene is more conspicuous than in the discontinuous three-color image, probably because of its relatively higher luminance. Also, the structure of the river and the road in the bridge scene are more apparent and the shading of red and blue in the lower intensity returns gives an illusion of three-dimensional viewing. The extraneous color boundaries are less apparent in the train scene and are much less objectionable than they were in the discontinuous three-color image.

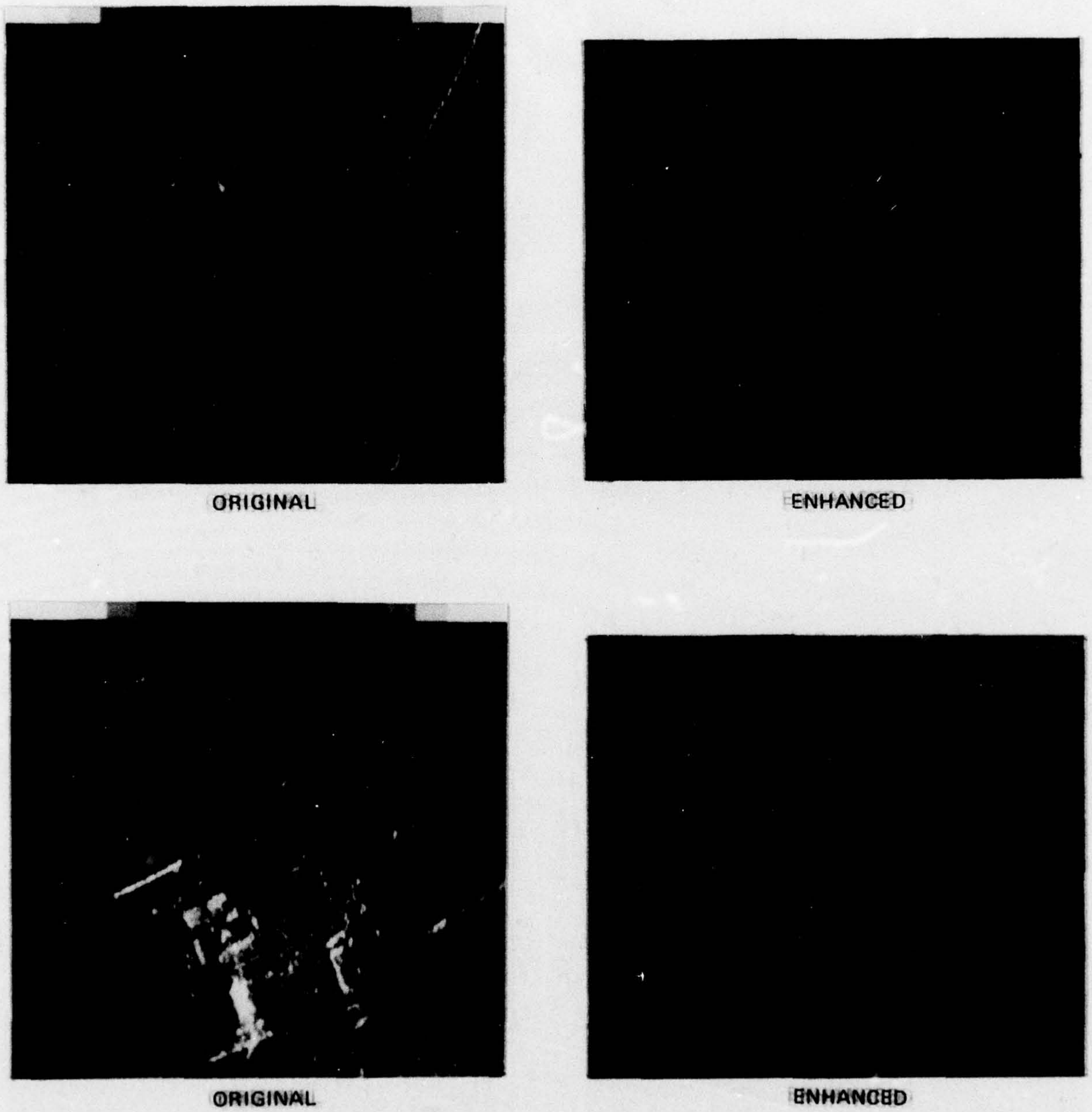


Figure 38. Three-color discontinuous intensity pseudocolor coding.

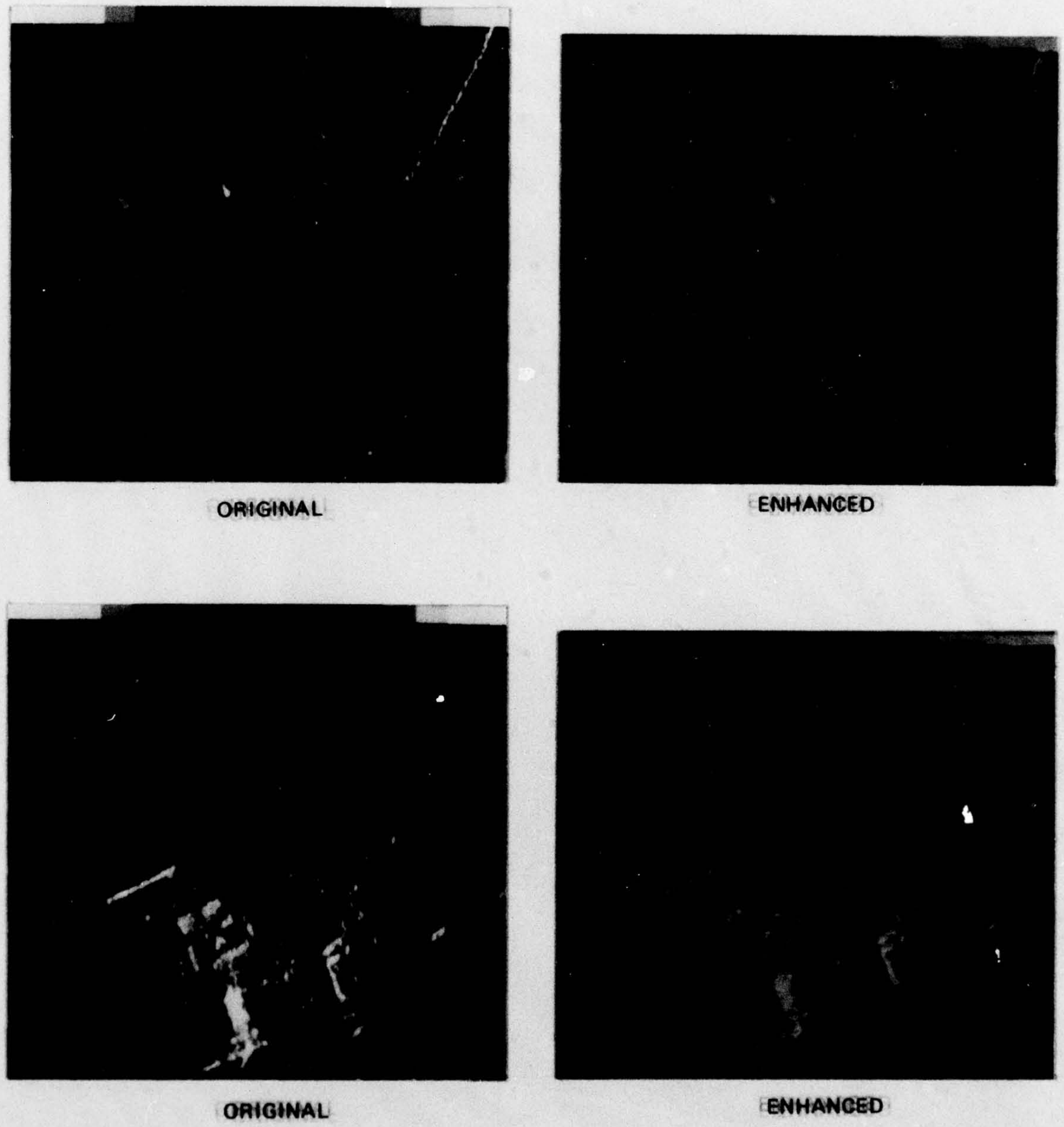


Figure 40. Three-color continuous intensity pseudocolor encoding.

Optimum placement of the color transition threshold levels is an important parameter in pseudocolor encoding. Conceptually, it would be desirable to set the threshold levels so that targets would be confined to a single color band and non-targets would be excluded from that color band. In practice, however, this does not appear to be achievable.

The pixel intensity histograms and cumulative intensity distributions shown in Figures 41, 42, and 43 for the bridge, train, and Barstow scenes indicate why.

The Barstow scene is a FLAMR SAR image of an Air Force test array of corner reflectors and various types of military vehicles arranged in different orientations located near Barstow, California. The various vehicle types and orientations allowed an analysis of the statistics of the radar returns from these vehicles. The mean radar return level from each type of vehicle is indicated in Figure 43, as well as the position of the mean return level of the entire Barstow scene, including the test array, part of the city of Barstow, and parts of the surrounding desert area.

The mean return (gray level) information ranged from a low of 89 for jeeps to a high of 108 for trucks; this range corresponds to the 83rd and 94th percentiles of the Barstow scene. Thus only 9 percent of the target image points in the Barstow scene are in this range. However, for the bridge scene, this same target image range corresponds to the 33rd to the 70th percentiles, containing 37 percent of the image points. For the train scene, the range corresponds to the 42nd to the 74th percentiles, containing 32 percent of the image points. Therefore, the range of mean return intensities of military vehicles can cover approximately one-third of the total number of points in realistic SAR images. When the variations in return intensities of the military vehicles are taken into account, the situation becomes worse.

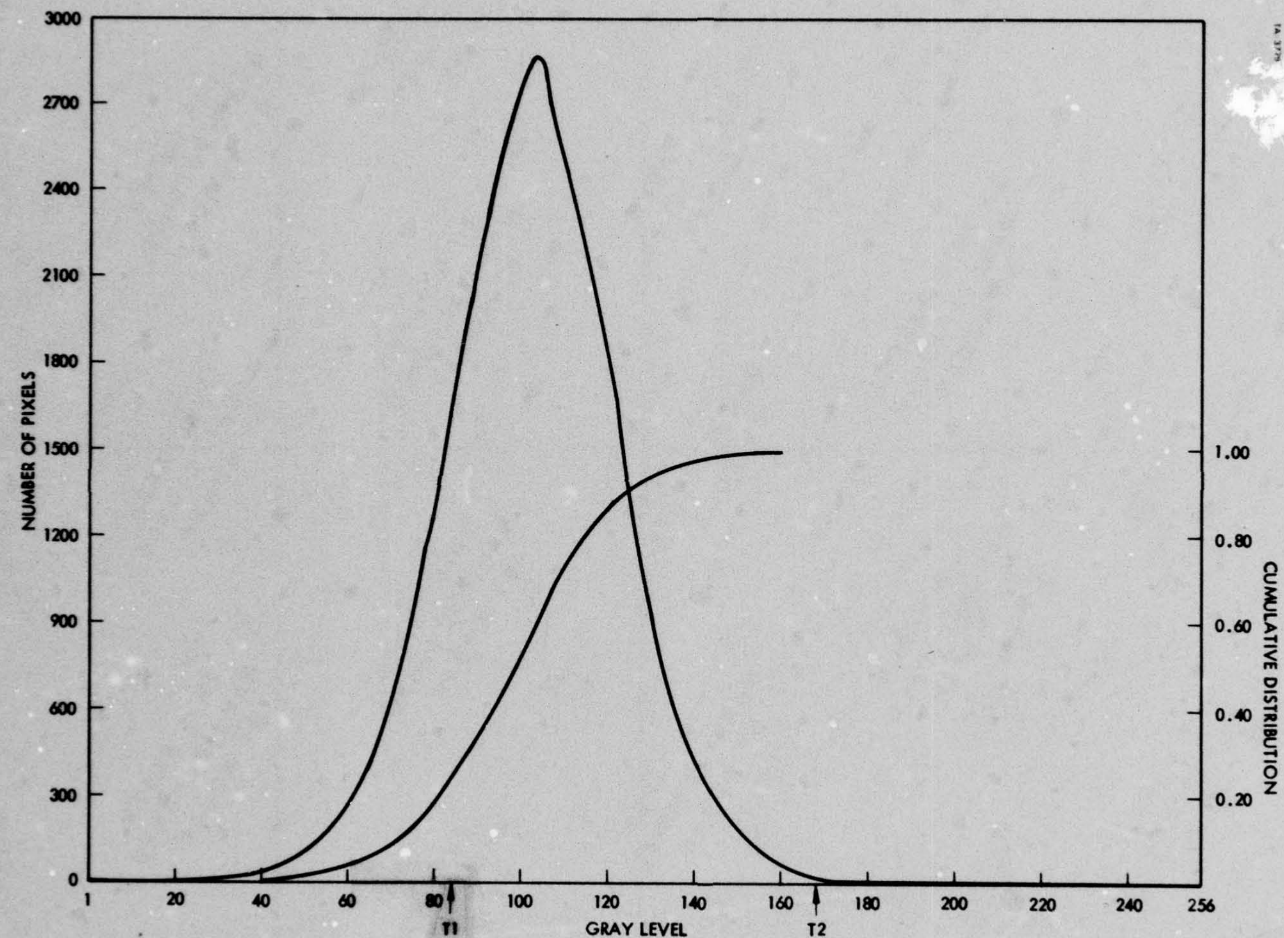


Figure 41. Bridge scene statistics.

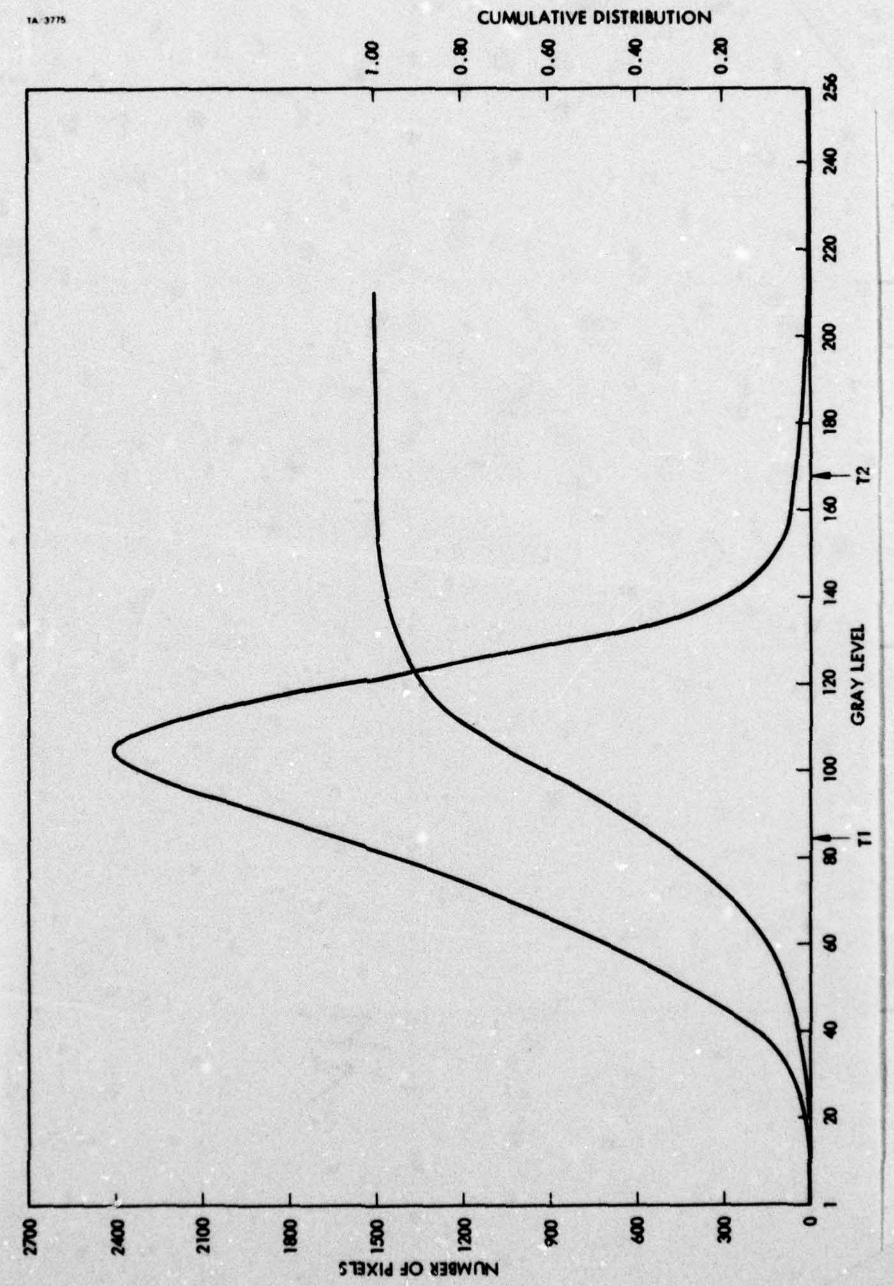


Figure 42. Train scene statistics.

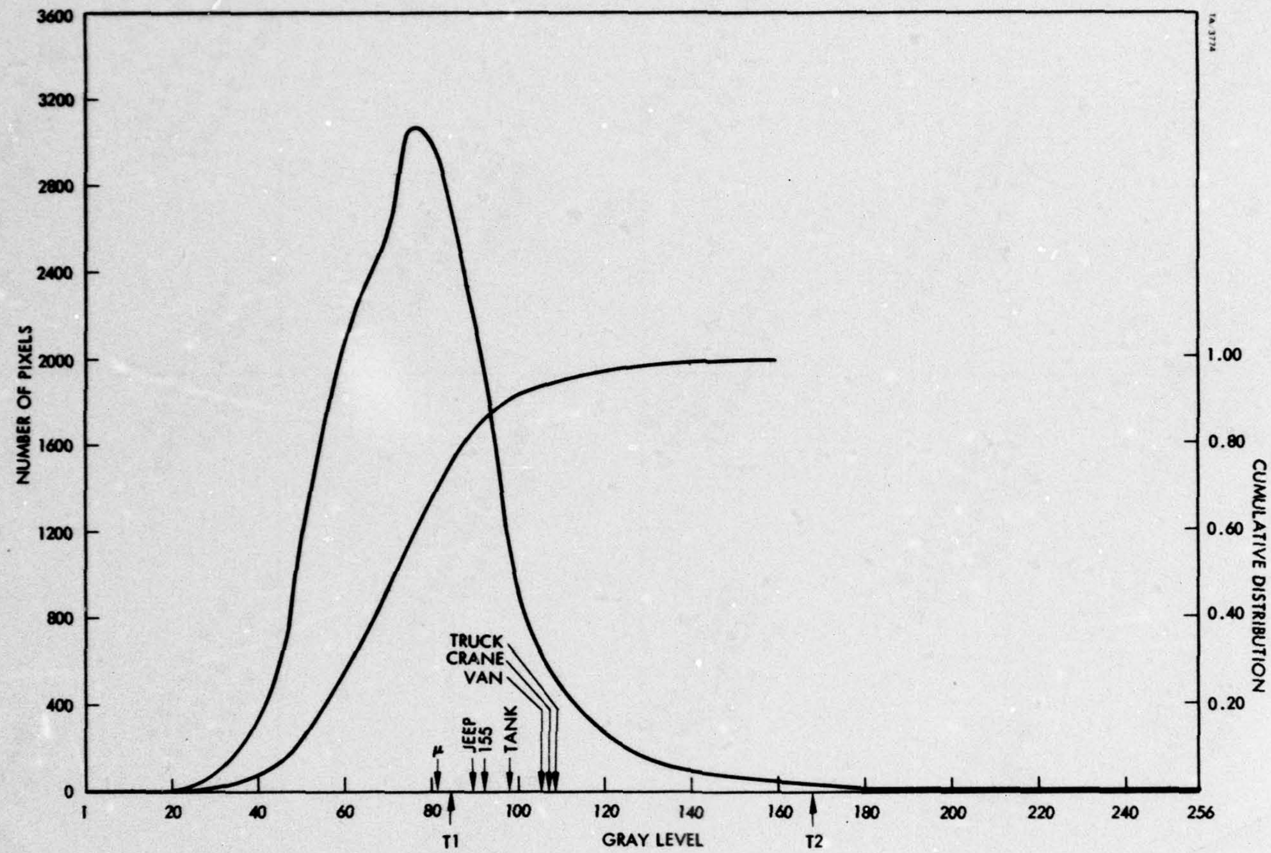


Figure 43. Train scene statistics.

Table 6 lists the range of return intensities of vehicles in the Barstow array by vehicle type. The overall range is from a low of 79 for jeeps to a high of 130 for cranes.

Table 6. Radar Return Intensity
Data for Military Vehicles

Vehicle	Return Intensity		
	Low	High	Mean
Tank	91	110	98
Truck	94	126	108
Van	86	129	105
Crane	97	130	107
Jeep	79	105	89
155	85	95	92

This corresponds to the 68th to the 98th percentiles in the Barstow scene, or 30 percent of the total number of image points. For the bridge scene, the target return range extends from the 17th to the 94th percentiles — 77 percent of the image points, and for the train scene, the extent of the overall range is from the 28th to the 95th percentiles — 67 percent of the image points. The obvious conclusion that can be drawn from these statistics is that no static assignment of color transition thresholds would isolate targets from non-targets. Even dynamically adjustable thresholds would be limited because of the large variation among tactical targets in a single scene.

Hue Coding and Saturation Coding

The one-, two-, and three-color techniques used color as a discrete indicator of gross intensity range, both independently and redundantly. The final category of pseudocolor encoding we investigated used continuous variation of the other properties of color besides luminance, namely, hue and saturation.

17
25

The first technique tried used hue encoding to represent radar return intensity, holding luminance and saturation constant. The idea was that blue would represent low return objects and as return intensity increased the corresponding hue would move through the greens and yellows to red, which would represent high return objects. Unfortunately, the standard red, green, and blue CRT phosphors have relative luminous efficiencies of 1.0000, 4.5907, and 0.0601, respectively (Wyszecki and Stiles, 1967). The desired constant luminance was therefore impossible to achieve with the available equipment. Because of this, the concept was modified to provide hue encoding at constant saturation and as nearly constant luminance as could be obtained.

The second technique tried used saturation coding to represent radar return intensity, holding hue and luminance constant. Again it was not possible to achieve constant luminance. The concept of saturation coding was that low intensity returns would be represented by white, i.e., unsaturated points, while higher intensity returns would be shown by points having a more saturated color.

Spectral colors with wavelengths between 480 and 640 nanometers were used for the hue encoding investigation. This was based on data (Kling and Riggs, 1971) indicating a just noticeable difference of 5 nanometers or less in wavelength throughout this band, insuring at least 32 distinguishable colors. The spectral areas below 480 nanometers and above 640 nanometers were discarded for two reasons: 1) the color discrimination capability of the eye falls off in these regions and 2) it is difficult to obtain a blue with high luminance and high saturation on a CRT display. The saturation level was chosen to be the maximum that could be recorded on film from the CRT display with color fidelity, bearing in mind the saturation loss effects of the film recording process and the nonlinear compensation for those effects. The saturation level used was 44 percent of the maximum saturation capability of the CRT. This resulted in an actual saturation, or color purity, of 37 percent for blue, 33 percent for green, and 43 percent for red, obtainable on the photographic film after compensation for CRT and film nonlinearities.

Figure 44 indicates the locus of the actual colors obtained in the hue encoding process on the 1960 CIE UCS chromaticity diagram. The gamut of CRT-producible colors restricted the degree of saturation and the spectral location of the endpoints of the band of colors used in the hue coding

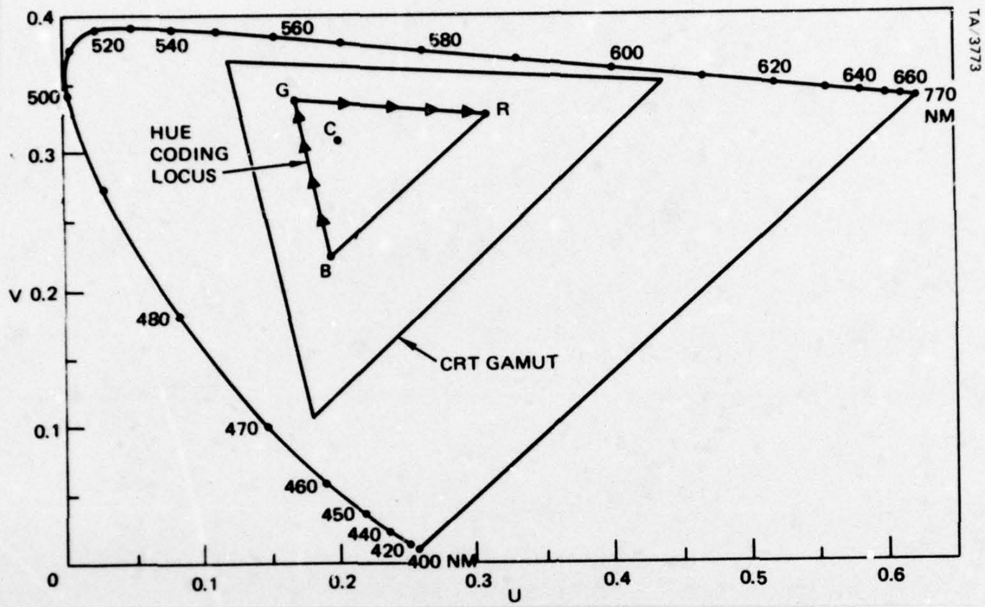


Figure 44. Locus of hue encoding colors on 1960 CIE UCS chromaticity diagram.

pseudocolor map. Thus the actual band of colors used was 460 to 620 nanometers. The relative luminances of the colors used in the hue coding map are shown in Figure 45.

Figure 46 shows two hue encoded pseudocolor SAR scenes. Both the bridge and the convoy in the bridge scene are quite apparent against their backgrounds of a different color. The high intensity structure of the bridge is also brought out more clearly by the red associated with very strong returns, compared to the unenhanced original. Although the river was made more distinct by the contrasting blue and yellow returns from the water and banks, the road appears more difficult to see against the background of mid-intensity returns than it was in the original. Similarly in the train scene, the segregation of high intensity returns into reds, oranges, and yellows reveals additional information in the train and power plant area that was not perceivable in the original. The train scene shows some effects of color edges

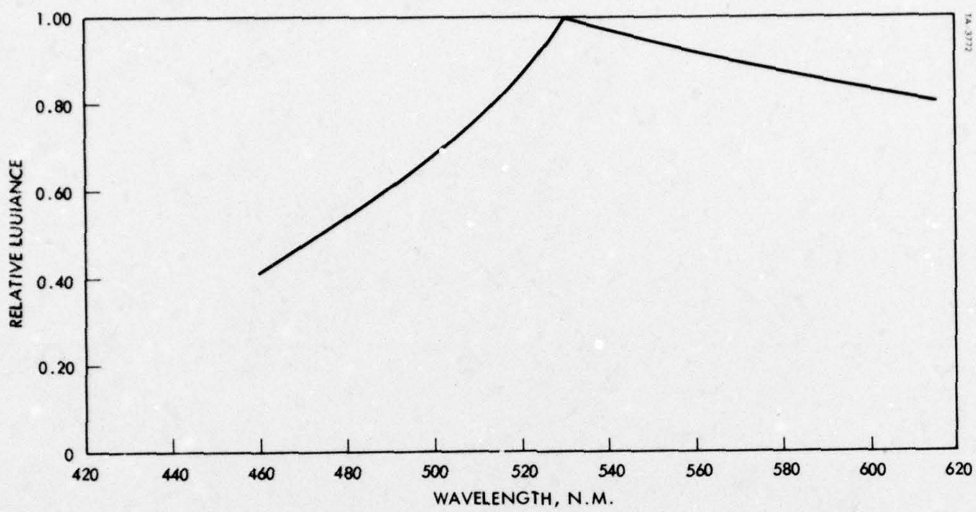


Figure 45. Relative luminance of spectral colors used in hue encoding.

tending to prevent, or make more difficult, the perception of continuous terrain. This might be minimized, as it was for the three-color map, by reducing the luminance of the colors representing the low intensity returns.

The hue encoded pseudocolor SAR images are "pretty". How much, if any, operator performance improvement might be realized is difficult to predict. Our judgment is that SAR hue encoded images would be most useful for intelligence extraction because of their potential for making small radar return intensity differences more perceptible.

Saturation coding was the last pseudocolor technique investigated. The three display primary colors — red, green, and blue — were evaluated, because they would be easiest to mechanize in an operational display system. Figure 47 shows the locus of saturation coding colors on the 1960 CIE UCS chromaticity diagram. The locus for red saturation is shown by the solid line; green and blue saturation are shown by the dotted lines. As mentioned previously, because of CRT restrictions and nonlinear film compensation, the maximum saturation level obtainable on film was 44 percent of the maximum

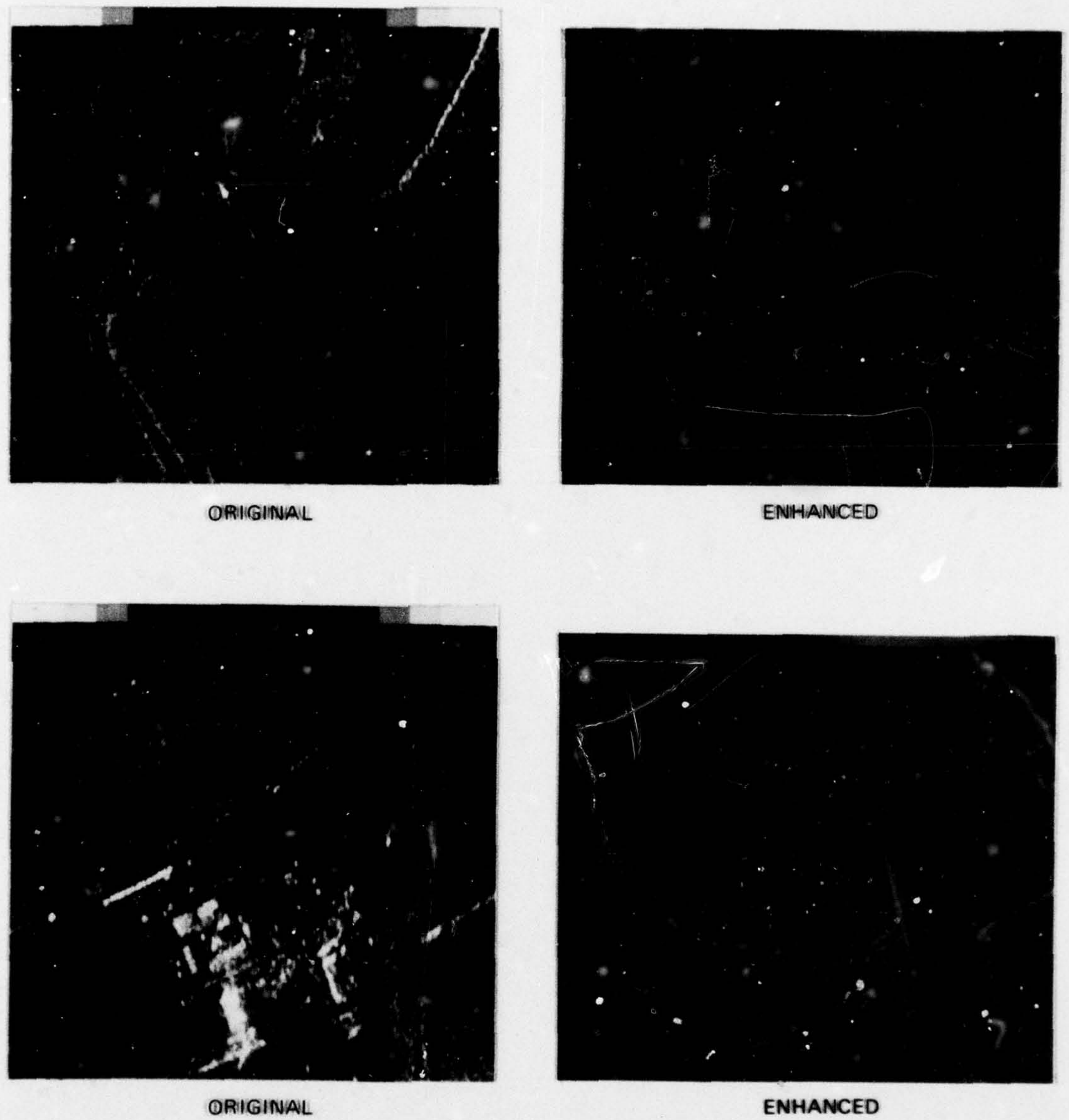


Figure 46. Hue encoding enhancement.

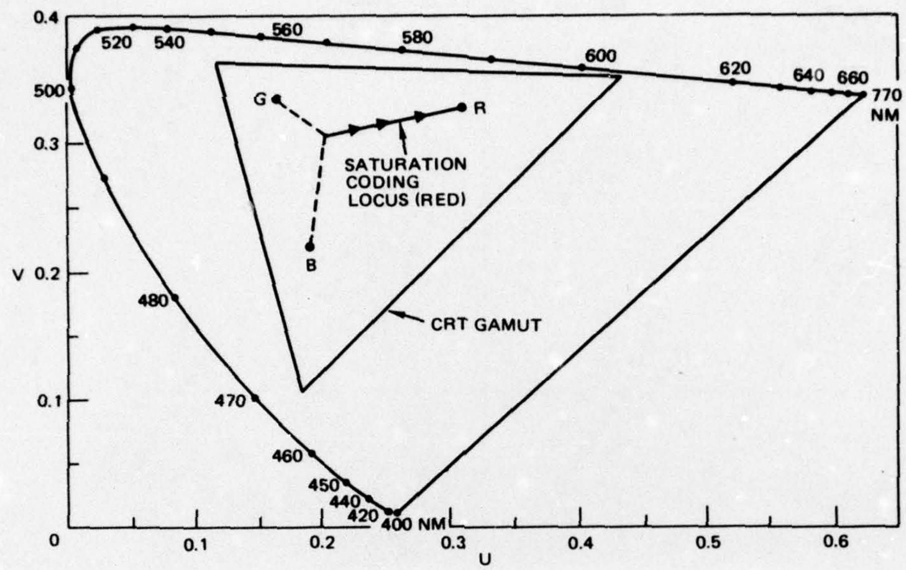


Figure 47. Locus of saturation coding colors on 1960 CIE UCS chromaticity diagram.

saturation capability of the CRT. This limited the maximum obtained saturation to 37 percent for blue, 33 percent for green, and 43 percent for red.

This limited dynamic range of saturation resulted in the number of discriminable colors being limited to 10 for green and 12 for red and blue. This is less than the 13 or 14 discriminable intensity levels in an image processed with the standard FLAMR gamma shaping function. Moreover, the 10 to 12 discriminable saturation level values were obtained by inspection of the color wedge at the top of each image. Within the scenes, only about five distinct saturation levels could be perceived in any of the hues. The difference in distinguishable saturation levels in the wedge and the scene was probably because areas shown at a given saturation level in the scene were smaller than those appearing in the color wedge, and sequential levels of saturation do not appear in corresponding spatial sequence in the image as they do in the color wedge. In any case, the results of the saturation coding were very disappointing. Some of the targets and scene features became

22
LJ

almost totally obscured, particularly in the green version. A suggestion was made that the sense of the coding be reversed, i.e., low intensity returns be represented by highly saturated color points and higher intensity returns be represented by less saturated points. Such a modification might produce some improvements, but it is believed that the restrictions on saturation imposed by the hardware account for most of the difficulties with this technique.

CONCLUSIONS AND RECOMMENDATIONS

The applicability of several forms of pseudocolor encoding to the display of synthetic aperture radar was examined from several standpoints: 1) as an independent discrete code to extend the dynamic range of information displayed quasi-continuously along another stimulus dimension, 2) as a redundant discrete code to emphasize and reinforce information presented along another stimulus dimension, and 3) as a new stimulus dimension to replace luminous intensity to display radar return intensity information.

The use of color as an independent coding dimension to extend the displayable SAR dynamic range seems to lead to ambiguity and difficulty of interpretation. Moreover, attempts to extend the amount of dynamic range information presented on a display beyond about 30 db are probably not worthwhile. There is evidence (Humes, Craig, Poplawski, Guerin, and Hershberger, 1974) obtained from a study performed using a laser display device, that target acquisition performance with 50 db of displayed dynamic range is no better than the 30 db of displayed dynamic range. This is consistent with the fact that the instantaneous perceivable dynamic range of the eye is no greater than 30 db (Kling and Riggs, 1971).

The interpretive difficulties associated with using color as an independent coding dimension are to a large extent eliminated by using color as a redundant coding dimension. The employment of color in this manner leads to more aesthetically pleasing presentations than the conventional intensity modulated monochrome display. Whether redundant color coding has any beneficial impact on the SAR operator target detection performance has not been assessed. The judgment of observers at Hughes is that the improvement in performance would not be sufficiently significant from an operational standpoint to justify the additional complexity and cost of a color display system, particularly in tactical aircraft.

23 LJ

The use of color as a stimulus dimension in place of intensity modulation may have a useful application in intelligence extraction and photogrammetry, particularly if the luminance profile of a hue encoded image could by dynamically adjusted by the human operator. It may also be desirable to dynamically vary the widths of the input intensity bands covered by colors. Flexible and dynamically variable pseudocolor mapping could provide more rapid and visually perceivable comparisons of return strengths from spatially separated portions of an image. The investigation of such techniques was beyond the scope of the present study; however, it represents an area in which further research might be pursued.

4.0 FEATURE ANALYTIC TECHNIQUES

INTRODUCTION AND RATIONALE

The general objective of feature analytic techniques is to employ automatic processing to search for areas within the image which possess a number of properties attributed to targets of a specific type, and to indicate to an operator the location(s) of areas having the properties of interest. Thus, the operator's search of the displayed scene and processing of fixated areas are directed toward portions of the image which have a high probability of containing a target. The notion is that a feature analytic device would result in cue directed operator visual search that would be more efficient than visual search resulting from the usual determinants of fixation sequences and durations.

Digital processing of SAR image data is not restricted to dynamic range limitations as are display devices and human perception. Machine processing of numerical SAR image data can be perfomed accurately and at high rates. While it might be argued that a human operator processes a large amount of image data in a relatively short time to construct a visual percept, and therefore achieves a fairly high throughput rate, the nature of such processing is not quantitative in an absolute sense. Construction of a detailed percept, such as is required for target detection or recognition generally will take significantly more time than that needed to obtain a general percept. The point is that automated machine screening of image segments in a fixed, structured sequence may occur rapidly enough to produce cuing information regarding areas containing targets. This screening should be faster and more reliable than the unaided operator. The reason is that the average amount of time spent in machine processing a given image segment would be several orders of magnitude smaller than the human observer. The processing time advantage would more than make up for the loss of efficiency associated with using a fixed, structured scan sequence.

Another reason to consider automatic search and cuing is reduction of the operator's workload and vigilance burden. Reduction of the operator's target acquisition task workload would allow him more time for other cockpit duties. Theoretical advances in pattern recognition and feature analysis, coupled with technological advances in electronic components and digital processing hardware, combine to make the feasibility of implementing feature analytic processors a reality.

BACKGROUND

The general problem of pattern recognition is one of classifying each member of a group of objects as belonging to one or another of several categories. The classification decision is based on: 1) a set of properties which characterize the elements of each category, and 2) the degree of resemblance between the properties of a given object and the characterizing properties of members of a given category. Thus all objects which are classified in a given category to some extent share the properties associated with that category. The nature of the characterizing properties depends on the nature of the objects of interest and the recognition requirements of the specific situation.

Applications in which pattern recognition techniques have been implemented in hardware with some degree of success are the reading of printed or typed characters, classification of electrocardiograms, and machines which play checkers or chess. Further potential applications of pattern recognition technology are reading handwritten characters, classification of X-ray images and other medical data, weather prediction, recognition of spoken information, and target detection and recognition in sensor imagery.

The property of feature invariance is a key to the success of any practical pattern recognition system, for in the real world variability will always be present. In the current SAR tactical target detection application, for example, it is desirable that the classification of an object as a tactical target be independent of its position within the field of view, independent of aspect angle, independent of background clutter, insensitive to incompleteness of pattern due to a missing element, insensitive to minor changes in pattern characteristics, and insensitive to changes in radar sensor parameters. The derivation or definition of features having some or all of these properties constitutes much of the art of pattern recognition, since there exists very little general theory to guide the feature selection process. For this reason the selection of features is often done empirically on an iterative basis.

The design and implementation of the decision device which performs the classification is, in concept, more straightforward than the design and implementation of the feature analyzer. Suppose that N features are measured from an input object, and their values are elements of an N-dimensional feature vector X. Then X is an element of an N-dimensional feature vector space Ω, and the classification problem is mathematically one of partitioning the feature space Ω into M mutually exclusive collectively exhaustive regions ω_i, i = 1, 2, . . . , M, where M is the number of categories into which the objects of interest are to be classified, and each ω_i corresponds to one particular category. A scalar discriminant function $D_j(x_1, x_2, \ldots, x_N) \equiv D_j(X)$ of the N feature variables can be defined for each region ω_j, j=1, 2, . . . , M, with the property that if the feature vector X lies in region ω_j, then $D_j(X) > D_i(X)$ for all values of i between 1 and M, except of course for i=j. Given the feature vector X and a set of M discriminant functions $D_j(X)$, j=1, 2, . . . , M, the mechanization of the decision device is fairly simple. The classifier computes the values of each of the $D_j(X)$, finds the maximum of the M resulting numbers, and assigns X to category k, if $D_k(X) = \max_{1 \leq j \leq M} D_j(X)$. Of course, many different functional forms can be selected for the $D_j(X)$.

No attempt is made here to summarize the vast body of pattern recognition literature. The interested reader is referred to the texts by Fu (1968) and by Duda and Hart (1973) for introductory material, and to the collection of papers edited by Mendel and Fu (1970) for more advanced material. In addition, the proceedings of the various conferences on pattern recognition are an important source of information, as well as the technical journals of, for example, the Computer Society or the Systems, Man, and Cybernetics Society, both published by the Institute of Electrical and Electronics Engineers (IEEE). For more applied information, the reports of industrial contractors are sometimes useful; although, they are sometimes difficult to obtain due to proprietary restrictions. Two recent examples which are available are Bullock (1976) and Geokezas (1976).

APPROACH

The critical element of a successful pattern recognition device is the definition and selection of feature measurements to be made. Consequently, the approach taken in this initial investigation of SAR tactical target pattern recognition techniques was to concentrate on feature definition. The major tool for performing this effort was the Scene Analysis Software (SAS) package developed, in part, at the Hughes Research Laboratories. This package consists of a number of mathematical operators which extract particular properties of images and a number of utility operators for performing image manipulations. The mode of operation was on-line interaction with non-real-time mathematical operators.

The effort consisted of applying various operators to images or portions of images. Individual operators were tried first; several operators were applied in sequence later to determine how the outputs of these operators varied with the content of the images to which they were applied. For example, a quantity of primary interest was the nature and amount of variation of an operator's output as a function of whether the image being operated upon contained a target or not. Output quantities which are sensitive to the presence or absence of a target by definition are good candidates for features to be extracted in an automated recognition device. Another quantity of interest was the relative computational speed of the various operators, particularly comparisons of computational speeds of operators whose outputs were essentially the same feature. For example, three versions of edge detection operators were evaluated in terms of processing speed and quality of edge detection performance.

Since the SAR imagery provided by the Air Force did not have good known examples of tactical targets, synthetic target signature complexes were created using the image manipulation capabilities of SAS package. In the process of creating these synthetic targets and imbedding them in the FLAMR imagery, care was taken not to introduce artificial properties which would make the synthetic targets have features not shared by real radar target signatures. In fact, one of the artificial targets was created from numerical data obtained by digitizing the actual radar return signature of a real target in a film transparency produced by an analog processed radar. The following describes in greater detail the targets which were imbedded in the FLAMR imagery and the nature of the operators applied to these images using the SAS capabilities. The results of each process are also described.

TARGETS

Four targets were used during the feature analysis experiment. These targets were developed and used at different times. The first two targets were simulated radar images of a mobile SAM site and a convoy. Later on during the program, the two more targets were added. These were a AAA site taken from actual analog processed radar returns and a section of the digitally processed Barstow array chosen to resemble a convoy.

Simulated SAM Site and Convoy

The simulated mobile SAM site and convoy were 8-bit 10 by 10 arrays. Before they were imbedded in FLAMR SAR images, a constant value was added to each pixel in the arrays so that the average values of the FLAMR images and the simulated target images matched.

AAA Site

The AAA site was composed of eight object returns from an analog processed radar scene arranged in a 30 by 30, 8-bit array.

Barstow Array

A 25 by 25 subsection of the Barstow array SAR image was extracted for use as the fourth target. The diagonal line of returns running from top left to bottom right represented the convoy.

SAR BACKGROUND IMAGE

The FLAMR SAR image used as the background image for the targets was the bridge scene used in the monochrome and pseudocolor enhancement work. A 128 by 128 subsection of the original 256 by 256 image was used to reduce processing time.

PROCESSING AND RESULTS

The ultimate goal of feature analysis is to distinguish between image areas containing targets and image areas which do not contain targets. Therefore, this initial effort examined the effect of several possible operators on the imbedded target images and image areas without targets to determine whether their output would be useful in a programmed decision tree that would cue human operators to probable target areas.

The operators used to process imagery in this study included: a statistics operator, three edge detection operators, a moment operator, and an operator for determining the centroid of an object. Not all four target area images were processed with each operator, but a sufficient number of examples were obtained to indicate the potential of each operator.

Statistics Operator

The statistics operator gave the minimum, maximum, area (in number of pixels), mean, and standard deviation of the processed image sub-section. Since the desire was to simulate as closely as possible the functions of an automated feature analyzer, the procedure was to partition the SAR image into sub-images. These sub-images were then processed using the statistics operator and the output examined for some criteria which could point to the presence of a possible target in the sub-area.

Nine sub-images were examined for each target area. Eight of these were background areas without targets, and one contained the entire target. an additional sub-image which contained only part of the target was processed for the SAM site. The processing was repeated on several of the sub-images with the intensities of the target returns reduced. This was accomplished by adding either 40 or 80 to the pixel value of the target returns, where 0 = white and 255 = black.

The results of the processing on the SAM site are given in Table 7. Sub-images 5 and 10 contained the target. Sub-image 5 contained the complete target while sub-image 10 contained the upper left corner of it. Table 8 lists the results of processing the convoy area. Again, sub-image 5 contained the target. The original convoy target area, without the convoy, is included at the bottom of the table.

Examination of the output of the statistics operator reveals two important points. In nearly all cases where the processing was performed in a sub-section without a target, there was a small standard deviation, and the mean was approximately halfway between the minimum and maximum, indicating a relatively uniform area. Where target sub-sections were processed, there was a larger standard deviation and the mean was biased toward the maximum value, indicating the presence of target returns. When the average modulation between target and background was reduced from 94 percent (the original) to 70 percent, the larger standard deviation and biased mean were still in evidence. The statistical information became ambiguous when average modulation was reduced to about 45 percent. Thus, there is information in the output of the statistical operator which could prove useful for pinpointing areas containing possible targets.

TABLE 7. RESULTS OF PROCESSING SAM SITE AND LOCAL AREAS

SUB-IMAGE	TOP LEFT COORDINATES X, Y	SIZE	MINIMUM			MAXIMUM			MEAN			STANDARD DEVIATION		
			ORIG.	+40	+80	ORIG.	+40	+80	ORIG.	+40	+80	ORIG.	+40	+80
1	11, 11	12 x 12	125			203			166			15		
2	23, 11	12 x 12	121			245			165			20		
3	35, 11	12 x 12	112			231			160			19		
4	11, 23	12 x 12	117			233			167			17		
5	23, 23	12 x 12	10	50	82	226	226	226	169	171	173	38	32	28
6	35, 23	12 x 12	108			234			154			22		
7	11, 35	12 x 12	110			245			167			20		
8	23, 35	12 x 12	122			249			171			18		
9	35, 35	12 x 12	105			226			157			23		
10	20, 20	12 x 12	18	58	82	226	226	226	168	169	170	31	27	24

TABLE 8. RESULTS OF PROCESSING CONVOY AND LOCAL AREAS

SUB-IMAGE	TOP LEFT COORDINATES X, Y	SIZE	MINIMUM ORIG.	MINIMUM +40	MINIMUM +80	MAXIMUM ORIG.	MAXIMUM +40	MAXIMUM +80	MEAN ORIG.	MEAN +40	MEAN +80	STANDARD DEVIATION ORIG.	STANDARD DEVIATION +40	STANDARD DEVIATION +80
1	66, 56	12 x 12	32			191			142			32		
2	78, 56	12 x 12	117			240			167			20		
3	90, 56	12 x 12	91			243			158			22		
4	66, 68	12 x 12	113			244			161			20		
5	78, 68	12 x 12	0	40		231	231		169	168		37	33	
6	90, 68	12 x 12	122			227			168			17		
7	66, 80	12 x 12	124			203			164			15		
8	78, 80	12 x 12	107			243			171			21		
9	90, 80	12 x 12	127			241			166			16		
Original Convoy Area	80, 70	10 x 10	0			195			169			42		

Edge Detectors

Three edge detection operators were available for edge detector processing. In order of increasing complexity they were 1) Roberts edge detector, 2) Sobel edge detector, and 3) Hueckel edge detector.

The Roberts operator works on an array of four pixels,

A	B
C	D

and calculates the edge at pixel $A = |A - D| + |B - C|$.
If this value is above a specified threshold, the Roberts operator replaces pixel A with a value of 1 and continues to the next pixel. The resulting output is binary where edge pixels have the value 1 (black) and all other pixels are white (0).

The Sobel operator is similar to the Roberts operator, but the array of pixels used in the calculation is a 3 x 3 array,

A	B	C
D	E	F
G	H	I

The edge at $E = \left| (A + 2B + C) - (G + 2H + I) \right| + \left| (A + 2D + G) - (C + 2F + I) \right|$.
Again, the resulting value is tested against a threshold, and a binary output is obtained.

The Hueckel operator works on a circular area around a pixel to determine if the pixel is on a boundary between two different zones within the circle. If it determines that the pixel is a boundary pixel, it outputs a short line at the pixel location indicating the boundary direction.

The three operators were first run on the 128 by 128 SAR image without imbedded targets for time comparisons. For the Roberts and Sobel operators, the threshold was set at a value of 150 and the spacing was 1 (a spacing of 1 meant that every pixel was considered). For the Hueckel operator, the spacing was also set so that every pixel was considered. The results were as follows:

Operator	Time, Seconds
Roberts	4.6
Sobel	12.9
Hueckel	59.4

Due to its lengthy processing time and more than necessary detail in the output image, the Hueckel operator was excluded from further consideration. The Roberts and Sobel operators were then tested on the SAM site and convoy targets. These results are shown in Figure 48. As can be seen, the Sobel operator outputs a double edge when it detects edge pixels, resulting in a misleading appearance of the target. Also, it detects more edges than are necessary for identification and takes more than twice as long as the Roberts operator. Thus, further evaluation of edge detection as a feature analysis operator used the Roberts operator.

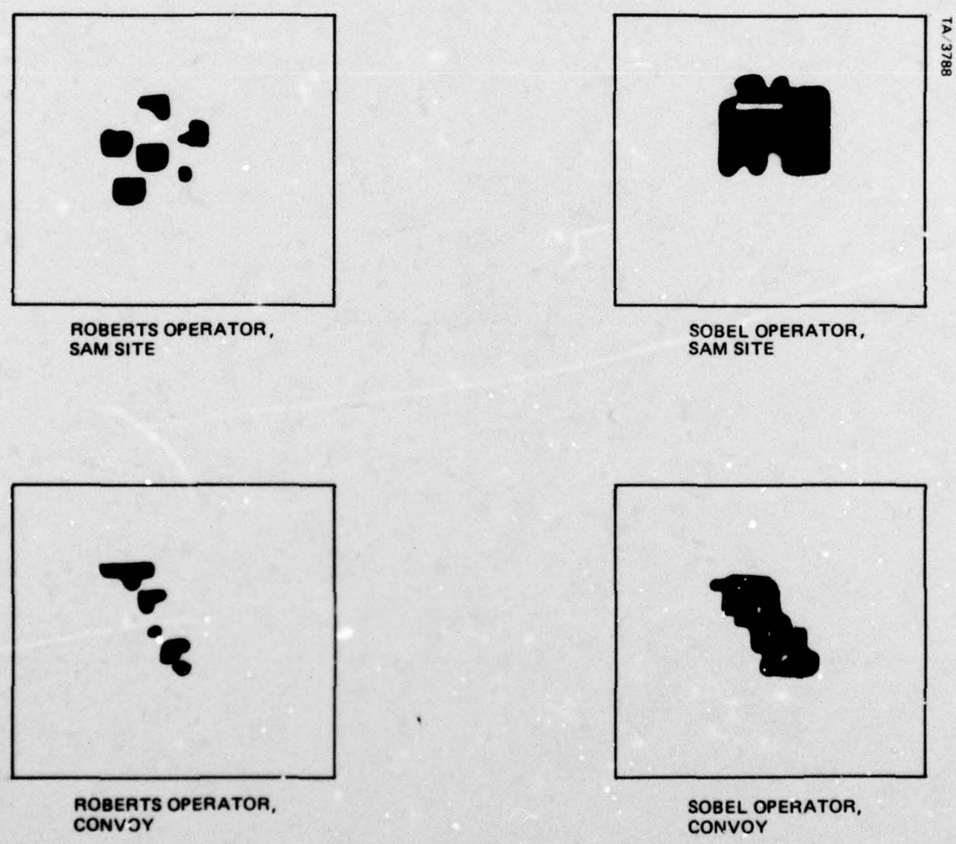


Figure 48. Roberts and Sobel edge detector results.

The prime motivation in investigating edge detection operators was to determine whether they could detect a sufficient number of edges to define an object which other operators could then further process. Analysis of the output data from further evaluation of the Roberts operator indicated more than adequate justification to consider additional operators to specify target shape and size information. The output image of the Roberts edge detector is well suited for such processing, because it is a binary image output from which nearly all information has been removed except that which is target related. Thus, it is likely that a boundary drawn to connect the edge pixels of the Roberts detected target will define the target limits. Once this boundary has been drawn, it is possible to calculate second-order moments for shape discrimination, area size, and location of the centroid of the target.

Moment Operator

Before processing by the moment operator, the edge points within the Roberts output images were connected. This was done by means of a utility operator which allowed replacement of a pixel value. In this case, pixels between those marked by the edge detector were changed from 0 to 1 to form the object boundary. Ultimately, this process would be automated.

The moment operator outputs several types of information. It finds the boundary of an object within an image, creates a silhouette image of the bounded object, and calculates several orders of spatial moments using first the boundary image and then the silhouette image. For this study, the primary interest lay in examining the second-order spatial moments as a means of distinguishing targets which are circular, ellipsoidal, or linear. The simulated SAM site and convoy were processed. Table 9 lists the second-order moments calculated from the boundary and silhouette images. What is important to note is the ratio μ_{20}/μ_{02}, where μ_{20} = moment about the horizontal and μ_{02} = moment about the vertical. This ratio remained of the same form despite variations in contrast. For the SAM site, a circular shape is suggested (the ratio is close to 1), while an elongated ellipse best fits the convoy (the vertical moment is considerably greater than the horizontal). In an automated feature analyzer, these moments would be compared to stored target reference models.

Table 9 also lists the output of an additional operator which calculated the centroid and number of target points of the silhouette picture from the

TABLE 9. RESULTS OF MOMENT OPERATOR

TARGET	Boundary					Silhouette		
	μ_{20}	μ_{02}	$\frac{\mu_{20}}{\mu_{02}}$	μ_{20}	μ_{02}	$\frac{\mu_{20}}{\mu_{02}}$	Centroid X, Y	Number of Target Points
SAM Site	4.79	6.12	.78	3.04	4.31	0.71	4, 5	43
SAM Site (+40)	5.61	6.79	.83	3.43	4.65	0.74	6, 7	48
SAM Site (+80)	4.65	5.99	.78	3.15	3.97	0.79	6, 7	39
Convoy	1.23	4.25	.29	1.56	5.04	0.31	5, 5	10
Convoy (+40)	2.89	7.21	.40	3.49	8.12	0.43	6, 6	15

moment operator. The centroid could be used to designate the center point of a target grouping for later reference. The number of target points would provide additional target classification information.

CONCLUSIONS AND RECOMMENDATIONS

From the evidence gathered by this study, the reality of a SAR tactical target feature analyzer looks promising. A possible approach, suggested by the initial work performed here, could be as follows.

An incoming image is quickly scanned sub-section by sub-section (the size being determined by the particular target grouping to be extracted). Basic statistics (maximum, minimum, mean, and standard deviation) are calculated for each area and compared against those of a stored reference model. Those areas which qualify as probable target areas are processed by an edge detector. The pixels identified as edge points are connected and second-order moments are calculated for the bounded object. These moments are also compared with those of the stored reference model, and if they match to within a specified range, the centroid of the object is determined, marked on the display, and stored for future reference.

Such a model is, of course, rather crude, but it forms a starting point for further experimentation and refinement. Some possible areas for future exploration could include: the addition of information such as road locations and the forward-edge-of-the-battle-area obtainable from a digital data base, the effects of thresholding sub-sections before the edge detection process to further reduce the chance of clutter returns being highlighted, and the feasibility of using a correlation process to extract target areas from an entire image which has been processed for edge detection.

There are many possibilities for branches of a feature analyzer's decision tree, and much work remains before the processes described here can be fully utilized. New targets need to be examined, the effects of contrast variations should be further explored, and an automated boundary drawing process needs to be developed, to name a few. The potential payoff, however, should be well worth the investment.

REFERENCES

Bullock, B. L., et al, "Finding Structure in Outdoor Scenes," Hughes Research Laboratories Report 498, July, 1976.

Burnham, R. W., R. M. Hanes, and C. J. Bartleson, "Color: A Guide to Basic Facts and Concepts," Wiley, New York, 1963.

Duda, R. O., and P. E. Hart, "Pattern Classification and Scene Analysis," Wiley, New York, 1973.

Fu, K. S., "Sequential Methods in Pattern Recognition and Machine Learning," Academic Press, New York, 1968.

Geokezas, M., "FLIR Image Analysis with the Autoscreener Computer Simulation," Honeywell Technical Report, February, 1976. AD A022755

Humes, J. M., D. W. Craig, T. C. Poplawski, D. F. Guerin, and M. L. Hershberger, "Synthetic Array Radar Operator Performance Evaluation Studies," AFAL-TR-74-158, November, 1974. (Confidential)

Humes, J. M., D. W. Craig, T. C. Poplawski, D. F. Guerin, and M. L. Hershberger, "Synthetic Array Radar Operator Performance Simulation Studies," AFAL-TR-74-63, August, 1974. (Confidential)

Judd, D. B., and G. W. Wyszecki, "Color in Business, Science, and Industry," second edition, Wiley, New York, 1963.

Ketcham, D. J., R. W. Lowe, and J. W. Weber, "Image Enhancement Techniques for Cockpit Displays," Phase II Final Technical Report, Hughes Report No. P75-497R, March, 1976.

Kling, J. W., and L. A. Riggs, editors, "Experimental Psychology," Holt, Rinehart, and Winston, Inc., New York, 1971.

Lamonica, G. L., "The Effects of Image Quality on Synthetic Aperture Radar Operator Performance," Volume 1, Target Recognition, AFAL-TR-74-333, March, 1975. (Confidential)

MacAdam, D. L., "Projective Transformations of ICI Color Specifications," J. Opt. Soc. Am., Vol. 27, p. 294, 1937.

Mendel, J. M., and K. S. Fu, editors, "Adaptive, Learning, and Pattern Recognition Systems," Academic Press, New York, 1970.

Nahi, N. E., "Estimation Theory and Applications," Wiley, New York, 1969.

North, D. O., "Analysis of the Factors Which Determine Signal/Noise Discrimination in Radar," RCA Tech. Rep. PTR-6-C, June, 1943. Also in Proc. IRE, Vol. 51, pp. 1016-1028, July, 1963.

Stockham, T. G., Jr., "Image Processing in the Context of a Visual Model," Proc. IEEE, Vol. 60, No. 7, July, 1972.

Van Trees, H. L., "Detection, Estimation, and Modulation Theory, Part I," Wiley, New York, 1968.

Wallis, R. H., "Film Recording of Digital Color Images," University of Southern California Image Processing Institute Report 570, May, 1975.

Wyszecki, G.W., and W.S. Stiles, "Color Science: Concepts and Methods, Quantitative Data and Formulas," Wiley, New York, 1967.

102

APPENDIX A. COLORIMETRY

The following discussion makes no claim of completeness or depth of treatment. For more detail, the reader should consult a reference book on colorimetry (Wyszeki and Stiles, 1967; Burnham, Hanes, and Bartleson, 1963; Judd and Wyszecki, 1963).

Fundamental to colorimetry and color matching is the so-called trichromatic generalization, which states that over a wide range of observation conditions many colors can be exactly matched or reproduced by additive mixtures of appropriate amounts of three fixed primary colors. The choice of the three primary colors is restricted only by the requirement that they be linearly independent, i.e., none of the primaries can be matched by a mixture of the other two. The trichromatic generalization requires that two colors which are initially matched must remain matched if each of their respective primary components is increased or reduced in amplitude by the same constant factor (linearity), and that if color A matches color B and color C matches color D, then the additive color mixture A+B must match the additive mixture C+D (additivity). Given a set of primary colors P_1, P_2, and P_3, any other color Q can be specified in the form $Q = \alpha_1 P_1 + \alpha_2 P_2 + \alpha_3 P_3$, where the α_i, called the tristimulus value of Q with respect to the primaries P_1, P_2, P_3, are the respective amounts of each primary required in an additive mixture which matches Q. Thus any color can be specified by its three tristimulus values with respect to some set of primaries.

Another way of specifying a color is through its spectral energy distribution $E(\lambda)$. This specifies the amount of radiant energy contained in a color Q as a function wavelength λ through the entire range of wavelengths in the visible spectrum. If the visible spectrum is divided into wavelength intervals of length $d\lambda$, then Q can be considered to be the sum of "elementary" colors $\bar{Q}(\lambda)$ whose spectral energy distributions are

$$\bar{E}(\lambda) = \begin{cases} E(\lambda), & \lambda \varepsilon \left[\lambda_i, \lambda_i + d\lambda\right], \\ O, & \text{otherwise} \end{cases}$$

where λ_i is the wavelength at the beginning of the $i^{\underline{th}}$ interval. Thus, $Q = \int \bar{Q}(\lambda)\, d\lambda$, where all integrals are over the entire range of wavelengths in the visible spectrum. If the tristimulus values of the colors $\bar{Q}(\lambda)$ are $T_1(\lambda)$, $T_2(\lambda)$, and $T_3(\lambda)$ with respect to some set of primaries $\bar{P}_1$, $\bar{P}_2$, and $\bar{P}_3$, and if T_1, T_2, and T_3 are the corresponding tristimulus values of Q with respect to the same set of primaries, $T_1 = \int T_1(\lambda)\, d\lambda$, $T_2 = \int T_2(\lambda)\, d\lambda$, and $T_3 = \int T_3(\lambda)\, d\lambda$ by the additivity property of the trichromatic generalization.

Rather than using the "elementary" colors $\bar{Q}(\lambda)$, whose properties depend on $E(\lambda)$, the spectral energy distribution of Q, it is more convenient to define a set of standard elementary colors $q(\lambda)$, namely the spectrum colors whose stimuli are all monochromatic and all contain the same constant amount of radiant flux. The spectral energy distribution $e(\lambda)$ corresponding to the standard elementary color $q(\lambda)$ is

$$e(\lambda) = \begin{cases} 1, & \lambda \varepsilon [\lambda_i, \lambda_i + d\lambda] \\ 0, & \text{otherwise} \end{cases}$$

where λ_i is as defined before. Then $\bar{E}(\lambda) = E(\lambda)\, e(\lambda)$, $\bar{Q}(\lambda) = E(\lambda)\, q(\lambda)$, and therefore $Q = \int \bar{Q}(\lambda)\, d\lambda = \int E(\lambda)\, q(\lambda)\, d\lambda$. Likewise, if $t_1(\lambda)$, $t_2(\lambda)$ and $t_3(\lambda)$ are the *tristimulus values of the colors* $q(\lambda)$ *with respect to the set* of primaries $\bar{P}_1$, $\bar{P}_2$, and $\bar{P}_3$, then $T_1(\lambda) = E(\lambda)\, t_1(\lambda)$, $T_2(\lambda) = E(\lambda)\, t_2(\lambda)$, and $T_3(\lambda) = E(\lambda)\, t_3(\lambda)$, so $T_1 = \int T_1(\lambda)\, d\lambda = \int E(\lambda)\, t_1(\lambda)\, d\lambda$, $T_2 = \int T_2(\lambda)\, d\lambda = \int E(\lambda)\, t_2(\lambda)\, d\lambda$, and $T_3 = \int T_3(\lambda)\, d\lambda = \int E(\lambda)\, t_3(\lambda)\, d\lambda$. By using a standard set of elementary colors $q(\lambda)$, the computation of tristimulus values T_i of a color Q from its spectral energy distribution $E(\lambda)$ is made simpler, because the tristimulus functions $T_i(\lambda)$, which vary with Q, need not be determined as an intermediate step. Instead, the functions $t_i(\lambda)$ can be found and used to compute the tristimulus values for an arbitrary spectral energy distribution $E(\lambda)$. The special functions $t_1(\lambda)$, $t_2(\lambda)$, and $t_3(\lambda)$ are called the color-matching functions with respect to the primaries $\bar{P}_1$, $\bar{P}_2$, and $\bar{P}_3$.

Two colors having spectral energy distributions $E_1(\lambda)$ and $E_2(\lambda)$ are color matched if and only if $T_1^{(1)} = \int E_1(\lambda)\, t_1(\lambda)\, d\lambda = \int E_2(\lambda)\, t_1(\lambda)\, d\lambda = T_1^{(2)}$, $T_2^{(1)} = \int E_1(\lambda)\, t_2(\lambda)\, d\lambda = \int E_2(\lambda)\, t_2(\lambda)\, d\lambda = T_2^{(2)}$, and $T_3^{(1)} = \int E_1(\lambda)\, T_3(\lambda)\, d\lambda = \int E_2(\lambda)\, t_3(\lambda)\, d\lambda = T_3^{(2)}$, i.e., their corresponding tristimulus values are all equal. Knowledge of $E(\lambda)$ uniquely determines the tristimules values T_i, but the reverse is not true. Hence, it is possible to obtain a color match in the sense defined above with many different spectral energy distributions. Two colors having the same tristimulus values but different spectral energy distributions are called metameric colors. Two colors having the same spectral energy distribution of necessity have the same tristimulus values and are called isomeric colors.

Associated with any set of tristimulus values, T_1, T_2, and T_3 is the set of chromaticity coordinates c_1, c_2, and c_3 defined by

$$c_1 = \frac{T_1}{T_1 + T_2 + T_3} \quad c_2 = \frac{T_2}{T_1 + T_2 + T_3} \quad c_3 = \frac{T_3}{T_1 + T_2 + T_3} .$$

Since $c_1 + c_2 + c_3 = 1$, it suffices to know the values of any two chromaticity coordinates. Because the relationships which define chromaticity coordinates are ratios, a constant scale factor applied to the tristimulus values will leave the chromaticity coordinates unchanged. Such tristimulus value scale factors correspond to variations in the luminance of the associated color, and therefore chromaticity coordinates contain no brightness information, only hue and saturation. Since chromaticity coordinates are defined from tristimulus values, the values of chromaticity coordinates depend on which set of color primaries is being used.

Several color primary systems have been developed and are commonly used. The first, and intuitively most appealing, is the RGB system defined in 1931 by the Commission Internationale de l'Eclairage (CIE). This system employs red, green, and blue monochromatic stimuli at 700.0, 546.1 and 435.8 nanometer wavelengths as primaries. The units are adjusted so that the chromaticity coordinates of the stimulus with equal radiant energy at all visible wavelengths are equal.

Figure 49 shows the (r, g) chromaticity diagram corresponding to the RGB primaries. If the tristimulus values of a color are R, G, and B in this system, then the chromaticity coordinates are

$$r = \frac{R}{R + G + B} \quad g = \frac{G}{R + G + B} \quad b = \frac{B}{R + G + B} .$$

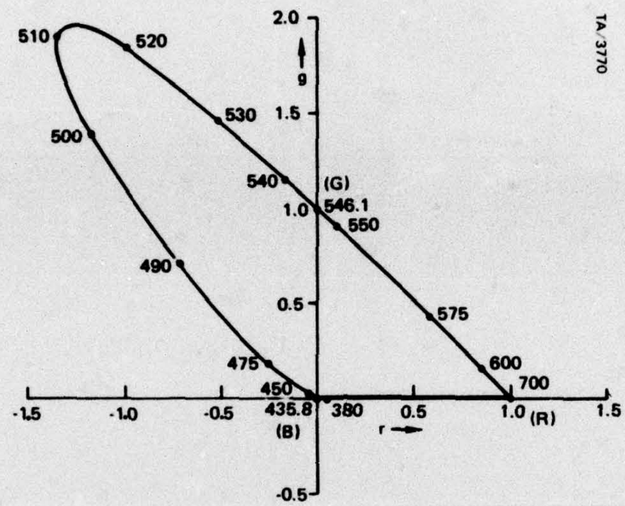


Figure 49. Chromaticity Diagram for 1939 CIE RGB Primaries.

Figure 49 is obtained by plotting r versus g for various colors. Note that b is redundant, since r + g + b = 1 by definition. The location of a color's chromaticity coordinates on the chromaticity diagram indicates its hue and saturation. For example, the chromaticity coordinate points of fully saturated spectral colors are shown by the corresponding wavelengths in nanometers along the periphery of the convex plane figure in the diagram.

As the saturation of a spectral color decreases, the corresponding chromaticity coordinate point moves from the edge of the diagram along a straight line toward the equal-energy stimulus point E in the diagram at coordinates (1/3, 1/3). Letting Q be any point on the chromaticity diagram and S the point on the edge of the diagram obtained by extending the line from E through Q until it intersects the edge, the saturation of chromaticity point Q is given by the ratio of the length of line segment EQ to the length of line segment ES.

The wavelength of the spectral color corresponding to the point S is called the dominant wavelength of the chromaticity point Q. If S lies on the r axis, no dominant wavelength is defined. In addition, there is defined a point $\overline{S}$, obtained by extending a line from Q through E until it intersects the edge of the diagram. If $\overline{S}$ corresponds to a spectral color, i. e., does not lie on the r axis, the wavelength of the spectral color corresponding to $\overline{S}$ is called the complimentary wavelength of the chromaticity point Q; otherwise no complimentary is defined. Every chromaticity point in the diagram has either a dominant wavelength, a complimentary wavelength, or in some cases, both.

The RGB system of primaries leads to negative chromaticity coordinates and tristimulus values for many points. This is inconvenient for computational work and instrumentation. For these reasons, the CIE in 1931 also defined the XYZ system of primaries. The XYZ primaries are imaginary, in the sense that they cannot be realized by actual color stimuli. However, the triangle formed by the chromaticity points corresponding to the imaginary primaries X, Y, and Z completely encloses the locus of chromaticity points corresponding to the spectral colors and the purple line connecting the two ends of the spectral locus. Hence, the chromaticity coordinates x, y, and z, defined in a manner analogous to the definition of r, g, and b, and the tristimulus values X, Y, and Z of any real color are all always non-negative. Figure 50 shows the (x, y) chromaticity diagram corresponding to the XYZ primaries.

Another desirable property of the XYZ system of primaries relates to matching visual stimuli in brightness. A visual stimulus with spectral energy distribution $E(\lambda)$ has a perceived brightness V given by $V = \int E(\lambda)V(\lambda)\,d\lambda$, where $V(\lambda)$ is a function of wavelength called the relative luminous efficiency function. The function $V(\lambda)$ indicates the variations in perceived brightness of monochromatic stimuli of unit radiant flux as a function of wavelength. Two stimuli with spectral energy distributions $E_1(\lambda)$ and $E_2(\lambda)$ are said to be matched in brightness if and only if $V_1 = \int E_1(\lambda)\,V(\lambda)\,d\lambda = \int E_2(\lambda)\,V(\lambda)\,d\lambda = V_2$. By design, the imaginary XYZ primaries were chosen so that the relative luminous efficiency function $V(\lambda)$ is the same as $y(\lambda)$, the color-matching function used to compute the Y tristimulus value from the spectral energy distribution. Thus, any two spectral energy distributions which have the same Y tristimulus value with respect to XYZ primaries will be perceived to be of equal brightness. Furthermore, any two spectral energy distributions which are color-matched with respect to the XYZ primaries will also be brightness matched.

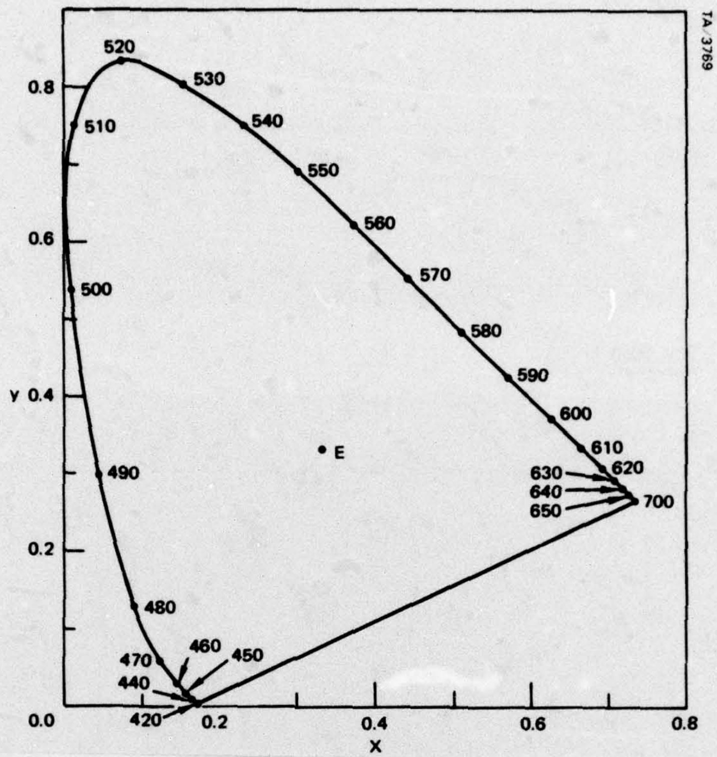


Figure 50. Chromaticity diagram for
1931 CIE XYZ primaries.

Despite the desirable properties designed into the XYZ system of primaries, there is one major disadvantage. Translation of two distinct chromaticity points by equal physical (Euclidean) distances on the (x, y) chromaticity diagram does not in general result in the perception of equal changes in the respective color stimuli. Although no simple transformation of the (x, y) chromaticity diagram will make equal translations correspond to equal perceptual changes in an exact sense, this property is so desirable that even a crude approximation based on a projective transformation would be useful. At least eight such transformations have been derived by various investigators based on different sets of experimental data (Wyszeki and Stiles, 1967). The chromaticity diagram which results from applying one of these transformations is called a uniform chromaticity scale (UCS) diagram. The chromaticity diagram resulting from the transformation derived by MacAdam (1937) was adopted by the CIE in 1960 as the standard UCS diagram and is referred to as the 1960 CIE UCS diagram.

The chromaticity coordinates in this system are generally denoted by u, v, and w, corresponding to the imaginary primaries U, V, and W. Figure 51 illustrates the 1960 CIE UCS diagram. The (u, v) chromaticity coordinates are obtained from the (x, y) coordinates by the relations

$$u = \frac{4x}{-2x + 12y + 3}$$

and

$$v = \frac{6y}{-2x + 12y + 3}$$

and as expected, $w = 1 - u - v$.

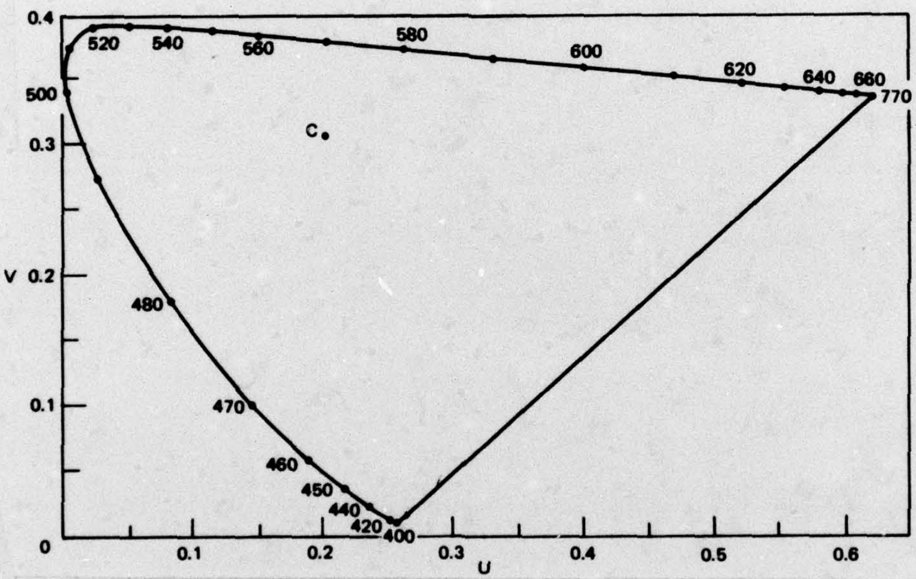


Figure 51. Chromaticity Diagram for 1960 CIE UCS System.